METHODOLOGY TO MODEL ACTIVITY PARTICIPATION USING LONGITUDINAL TRAVEL VARIABILITY AND SPATIAL EXTENT OF ACTIVITY

A Dissertation
Presented to
The Academic Faculty

by

Vetri Venthan Elango

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy in the
School of Civil and Environmental Engineering

Georgia Institute of Technology
December 2014

METHODOLOGY TO MODEL ACTIVITY PARTICIPATION USING LONGITUDINAL TRAVEL VARIABILITY AND SPATIAL EXTENT OF ACTIVITY

Approved by:

Dr. Randall Guensler, Advisor
School of Civil & Environmental Engineering
Georgia Institute of Technology

Dr. Michael Meyer
Parsons Brinckerhoff

Dr. Michael Hunter
School of Civil & Environmental Engineering
Georgia Institute of Technology

Dr. Michael Rodgers
School of Civil and Environmental Engineering
Georgia Institute of Technology

Dr. Catherine Ross
College of Architecture
Georgia Institute of Technology

Date Approved: August 27, 2014

To My Family

ACKNOWLEDGEMENTS

I wish to thank Dr. Randall Guensler for being my advisor and guiding me through this dissertation effort. I would like to thank Dr. Guensler for not only teaching me research skills but also for helping me learn to provide attention to details without losing focus on the big picture. I want to thank my committee members, Dr. Michael Meyer, Dr. Michael Hunter, Dr. Michael Rodgers, and Dr. Catherine Ross, for providing valuable feedback.

I would like to thank all my colleagues from the Drive lab who provided an intellectual and collegial environment to thrive. I would like to especially thank my colleagues Dr. Hainan Li, Dr. Angshuman Guin, Dr. Wonho Suh, and Dr. Yanzhi Xu with whom I had to the opportunity to work closely. The discussions during the team work were immensely valuable in developing my research skills.

Last but not the least, I wish to thank all my family and friends for supporting me through this dissertation effort. I wish to thank my father and mother for their sacrifices, without which I could not have accomplished this dissertation. I wish to thank my sister for her support and encouragement. I wish to thank my wife for being there through the ups and downs, believing in me and being a great supporter. I wish to thank my daughter for being the greatest joy in my life and providing me perspective on life. I dedicate this dissertation effort to my family.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

SUMMARY

Macroscopic changes in the urban environment and in the built transportation infrastructure, as well as changes in household demographics and socio-economics, can lead to spatio-temporal variations in household travel patterns and therefore regional travel demand. Dynamics in travel behavior may also simply arise from the randomness associated with values, perceptions, attitudes, needs, preferences and decision-making process of the individual travelers. Most urban travel behavior models and analysis seek to explain variations in travel behavior in terms of characteristics of the individuals and their environment. Spatial extents and temporal variation in an individual's travel pattern may represent a measure of the individual's spatial appetite for activity and the variability-seeking nature on his/her travel behavior. The objective of this dissertation effort is to develop a methodology to predict activity participation using revealed spatial extents and temporal variability as variables that represent the spatial appetite and variability-seeking nature associated with individual household. Activity participation is defined as a set of activities in which an individual or household takes part, to satisfy the sustenance, maintenance and discretionary needs of the household. To accomplish the goals of the dissertation, longitudinal travel data collected from the Commute Atlanta Study are used. The raw Global Positioning Systems (GPS) data are processed to summarize trip data by household travel day and individual travel day data. A methodology was developed to automatically identify the activity at the end of each trip. Methods were then developed to estimate travel behavior variability that can represent the variability-seeking nature of the individual. Existing methods to estimate activity space were reviewed and a new Modified Kernel Density area method was developed to address issues with current methods. Finally

activity participation models using structural equation modeling methods were developed and the effects of the variability-seeking nature and spatial extent of activities were applied to the models. The variability-seeking nature was presented in the activity participation model as a latent variable with coefficient of variation of trips and distance as indicator variables. The dissertation research found that inclusion of activity space variables can improve the activity participation modeling process to better explain travel behavior.

CHAPTER 1

INTRODUCTION

The Inter-modal Surface Transportation Efficiency Act (ISTEA), 1991 and the Clean Air Act Amendments (CAAA), 1990, place emphasis on accurately predicting changes in travel behavior over time. Behavioral changes typically result from changes in the transportation environment associated with new transportation policies and advances in technology [1]. Changes in travel behavior can also occur due to changes in the economy, land use development and demographic changes in the household. The emphasis on transportation planning has shifted from capacity expansion (building new infrastructure) to the formulation of transportation policies that effectively manage travel demand, which necessitates better understanding of travel behavior [2]. The ISTEA and CAAA reflect this shift in policy towards capacity management. Historically, travel behavior models were able to use average travel behavior to identify corridors for capacity expansion. Capacity management, on the other hand, requires better understanding of travel behavior to create policies that would reduce demand and allocate resources more efficiently.

Travel demand models and travel behavior models help to evaluate the impact of transportation plans, policies, and programs on travel behavior such as changes in the number of trips, route choice and time of day of the trips. The quantity and quality of the travel data used to build these models determines the quality of the models. Traditional travel diary surveys collect one or two days of travel data for each household. Hence, traditional travel diary survey methods collect cross-sectional data designed to represent a snapshot of travel behavior in time. Such cross-sectional travel data do a decent job of capturing average household travel behavior within specific population sectors when conducted in large numbers using proper random-stratified sampling techniques. However, these traditional surveys do not provide sufficient data for undertaking more detailed behavioral analysis at the disaggregate level [1, 3]. Multi-day travel data collected over longer periods in longitudinal studies can provide valuable information for analysis and estimating the microscopic level changes that occur at the household or the person level.

Macroscopic level changes in the urban environment and transportation system such as changes in transportation policy or new infrastructure, as well as microscopic level changes at the household level in demographic or socio-economic structure cause variations in travel patterns. Dynamics in travel behavior may also arise from the randomness associated with the values, perceptions, attitudes, needs, preferences and decision-making process of the individual travelers. Most urban travel behavior models and analyses try to explain variations in travel behavior in terms of characteristics of the individuals and their environment [4]. However, socio-demographic characteristics of the individuals do not represent the individual's adventure seeking nature, and the environmental constraints that influence the individual's travel behavior.

The need for realistic representation of travel behavior in modeling has led to the realization that traditional statistically oriented trip based modeling has to be replaced with a more behaviorally oriented activity based modeling approach [5]. Activity participation is defined as a set of activities that an individual or household takes part, in order to satisfy the sustenance, maintenance and discretionary needs of the household.

Activity Based Models in Practice

Metropolitan planning organizations (MPO) across the country have started to embrace activity based approaches to travel demand modeling. Planners typically simulate the performance of transportation networks in the future based on predicted population growth, predicted land use changes and impact of policies and changes to infrastructure. In the activity based modeling framework, activity participation instead of trip generation is modeled. The activity participation model leads to allocation of time to different activity types and activity locations by each member of the simulated population. The distribution of the activity locations across space necessitates the individuals to tour these locations. The tour generation is simulated based on the activity participation and the tours are assigned to the transportation networks. The assigned trips to the transportation infrastructure is then evaluated for performance measures.

The Atlanta Regional Commission started implementing a dual-track method in 2010 to maintain a four-step traditional travel demand model and implemen an activity based model using the Coordinated Travel Regional Activity Based Modeling Platform [6, 7]. The activity based model design combined advanced modeling tools, such as discrete choice forms, activity duration models, time-use models etc., to ensure behavioral realism, replication of the observed activity travel patterns, and ensure model sensitivity to key projects and policies [7]. The activity based modeling framework involved five steps. The first step was the synthesis of the population for the Atlanta planning region. The decision-makers in the model, households and persons, are synthesized for each simulation year based on census data. This was followed by the long term choices that relate to the usual workplace and school for individuals when applicable. The daily activity pattern of each household member is modeled in the next step. The activity patterns include mandatory, non-mandatory and in-home activities. Tour frequency, and choice of destination by time of day are modeled for each household member to satisfy the individual's activity. The fourth step involves modeling the tour-level details of mode, number of intermediate stops and location of stops. The fifth step assigns the trips to the mode and the trips are assigned to the highway and transit networks based on the mode [7].

The San Diego Association of Governments (SANDAG), a regional planning agency is in the early stages of transitioning their transportation demand modeling from the traditional four step trip based process to the activity based modeling framework. SANDAG is proposing to use a six step framework to develop their regional transportation planning based on activity based model [8]. The first step will create the synthetic population for the planning region based on census and land use models. The synthetic population will represent the decision makers whose travel choices will be input to the activity based model. The second step will assign the work and school location for each of the decision makers in the synthetic population if applicable. The third step will predict the mobility characteristics of individuals and households such as vehicle ownership, and parking costs. The accessibility of household locations will also be predicted as part of the mobility characteristics. A daily activity pattern will be scheduled for each individual in the synthetic population as part of the fourth step. The model schedules each tour type by predicting how many tours, type of tour (mandatory, non-mandatory or joint), participants, departure and arrival time and locations. The fifth step will predict the characteristics of each tour, such as primary mode, number of stops, location of stops, and departure time from stops. In the last step, the predicted tours will be aggregated, and assigned to the zone to zone matrix of origins and destinations. As part of this process trips are assigned to the transportation networks and the performance of the transportation network can be evaluated [8].

Activity participation modeling, which results in allocation of duration to each activity type, is one of the first steps in the activity based modeling framework. Activity participation models like other travel behavior models use socio-demographic characteristics as variables and do not have variables to reflect the household's preferences and constraints. The household's preferences and constraints contribute to the error in the activity based models just as in the traditional modeling techniques.

Dissertation Objectives

While the socio-demographic variables represent the opportunity and necessity to make trips, the socio-demographic variables do not reflect the specific preferences and constraints of a household. The objective of this dissertation is to develop a methodology to create activity participation models that incorporate spatial and temporal variability in addition to the socio-demographic variables. The underlying assumption in using spatial and temporal variability is that the spatial extent and the travel variability will help us capture individual's travel preferences, individual environmental constraints, and differences in individual random behavior.

To develop enhanced activity participation models, this dissertation proposes to convert longitudinal Geographic Positioning Systems (GPS) data from the Commute Atlanta Study into useful trip data, identify the activities under taken by the participants, quantify the travel variability and spatial extent of activities and use the new variables in building activity participation models. The first step was to process the passively collected GPS data from the Commute Atlanta study into a useful format. This involved doing quality assurance quality control on the raw GPS data, processing them to routes, and identifying trip chains. The next step would be to identify activities participated for the passively collected longitudinal trip data from the Commute Atlanta study. The third step was to evaluate the quantitative variability in the vehicle trips over time for each household. Estimation of activity space for each household was the next step in this process. The final step assessed the integration of activity space and travel variability as variables that represent the household's variability-seeking nature and appetite for spatial extent into activity participation models. The above steps will help in making passively collected longitudinal GPS data useful for modeling and estimate the travel behavior variability and spatial extent that can be used to enhance activity participation models. The dissertation will explore whether the enhanced activity participation models are better at explaining travel behavior than just using socio-demographic variables.

Scope

This dissertation focuses on processing passively-collected longitudinal GPS travel data to make the data in a useful format for activity based modeling and to develop an activity participation model using such data. The methodologies developed in this dissertation will be applicable to data collected from similar instrumented vehicle studies in major urban areas. However, to apply the developed methodologies on data collected using other technologies or from other cities, the methodologies will need to be suitably changed to accommodate the changes in assumptions. The assumptions used in developing the methodologies described in this dissertation will be stated in detail as part of the methodology development. The dissertation also acknowledges that transitioning the methodologies developed, to another region or dataset will involve significant re-evaluation of the assumptions and modifying the methodologies suitably.

The data employed in this research effort were derived from the Commute Atlanta study; in which instrumented vehicles were used to collect household travel data. Travel via other modes (walking, bicycling, transit, etc.) were not monitored in the study. Hence, this research assumes that the vehicle travel data are comprehensive for each household, and assumes that travel by non-monitored transportation modes are relatively insignificant with respect to regional travel behavior model development and application. This leads to the assumption that all trips were made by vehicles and walking mode component of each trip was less than a quarter of mile. The activity participation model developed in this effort is assumed to reflect all activities using vehicles. Vehicle trips constitutes 87% of trips in Atlanta where 94% of households have access to at least one vehicle [9, 10]. Households without auto access make most of the non-auto trips in Atlanta making the share of auto trips by households with auto much higher than 87%. From a planning perspective walk trips and bike trips are essential in assessing transit needs, sidewalks and bike routes. Non-auto trips do not significantly affect the ability of the four step transportation planning process in identifying new highway and arterial infrastructure needs. Future travel diary studies will involve collection of data via all modes by using devices such as smartphones that can be easily carried by the participant with ease. The methods developed in this dissertation are directly extensible to the data that will include transit and other alternative transit modes.

Research Methodology

To achieve the objectives of this dissertation, the research methodology includes five main tasks. The literature review seeks to provide an understanding of the existing methods and research findings in current literature on travel surveys, trip purpose identification, travel behavior variability, activity space extents and activity participation modeling. The second step is to process the raw instrumented vehicle data from Commute Atlanta into a useful format. The next task is to develop and apply methods to identify the activity purpose of each trip for passively collected GPS data. The fourth task is to estimate the temporal variability in the travel behavior of each households. Following that, methods to estimate the activity space for each household are developed and applied. The final step is to develop an activity participation model using the variability measures and spatial extent along with the socio-demographic variables.

Literature Survey

The literature review includes various topics relevant to the dissertation including:

- Travel Data Collection methodologies
- Activity Purpose Identification
- Variability in Travel
- Activity Space Estimation
- Structural Equation Modeling
- Activity Participation Models

Data Collection and Processing

The raw data from the instrumented vehicle fleet contains the trip information of the participating vehicles. Data from surveys on household demographics are stored in a relational database. Periodic surveys were undertaken to update the changes in the household and the database schema reflects the longitudinal nature of the survey [11]. The raw instrumented data are checked for quality control and quality assurance. The trip data are then route processed to identify the network paths. The raw trip data files include all GPS records from engine-start to engine-off. These files need to be processed to identify trip chains. Finally, the trip data and the socio-demographic variables need to be integrated for analysis. The integrated trip data are aggregated at the household day level to create a household summary of travel.

Activity Identification

The activity identification for passively collected instrumented-vehicle data is critical to make the data useful for travel behavior modeling. The disadvantage of passively collected data is that you have to surmise the activity undertaken at the destination. The relational database in the Commute Atlanta Study includes home, work, and school locations for each participating household. However, if a location changes and the household does not report the change to the study immediately, the locations in the database maybe incorrect until the household reports the change in a quarterly follow-up survey. Therefore, the activity identification involves the verification of known locations in the database, and identifying any changes to home, work or school locations. After identifying the known locations, the activity at the unknown locations can be identified using land use parcel data and geo-fencing techniques [12].

With the above limitations of passively collected travel data, the dissertation focuses on the following trip purposes: In-Home Activities, Out of Home Work Activities, Out-of-Home Sustenance Activities, and Out-of-Home Recreation Activities. If a location has more than one activity type possible, then it is classified as a Potential Multi-Purpose Activity.

The most frequently visited location for all the vehicles in a household is the home location and the last trip of the day usually ends at the home location. After the home location, the next most frequently visited is the work location. The work activity usually follows the same temporal cycle between weeks. The analytical methods will include the creation of a household table of frequently visited locations by time of day and day of week. This table helps to verify and correct the locations in the database.

The end location of each trip is then matched to the home, work and school locations to identify these activities. The next step is to compare the remaining trip-end locations with land use database to identify the activities associated with the locations. The data processing methods will assign the land use within a quarter-mile of the vehicle parked location for these trip end locations.

Variability in the Number of Trips

The next step is to study the temporal variability of the travel behavior and estimate those measures to use in Activity Participation modeling. The dissertation uses the trip data summary created to estimate the variability in the number of trips. The study evaluates different measures of variability such as the standard deviation, and coefficient of variation of number of trips, and distance for each household month. This effort analyzes the above measures of variability for the potential to accurately capture travel behavior variations and selects the appropriate measure to use in activity participation modeling.

Estimation of Activity Space

Estimating the spatial extent of activity is the next task. The study estimates activity space of all activities by the household in a month. The dissertation explores confidence ellipses, and Kernel Density methods to estimate these activity spaces. The dissertation develops a new methodology, Modified Kernel Density method, to estimate activity space to address the shortcomings of the above methodologies.

Activity Participation Model

To model activity participation, the dissertation explores the use of structural equation models. The activity space and travel variability along with socio-demographic variables will be part of the household variables that will explain activity participation and trip generation.

After identifying models that need to be created, the dissertation builds the activity participation model using the Commute Atlanta data. The results of the models helps to understand the significance of using travel variability and activity spaces as variables that explain the travel variability-seeking nature and appetite for spatial extent of the individual households. The dissertation will compare the model results with a model that only uses socio-demographic variables.

Dissertation Outline

The second chapter summarizes the review of the current literature on the various topics relevant to this research. The third chapter presents the data collection methods, and data processing techniques. Chapter three also addresses the data quality assurance and quality control methods used. The fourth chapter outlines the methods used to identify the activity purpose for each trip. The fifth chapter applies this methodology to a case study and evaluates the performance of the methods. The sixth chapter explores the variability measures that can help in capturing travel behavior variability and describes the methodology to estimate the variability measure. Chapter seven presents the results of the travel variability measures and explores the usefulness of each measure in accurately portraying the variability-seeking nature of the households. The eight chapter evaluates the current methods to estimate activate space and develops a new method to calculate activity space. The ninth chapter studies the estimated activity space in relation to the socio-demographic variables. The tenth chapter reports the activity participation model development using structural equations modeling. Chapter 11 presents activity participation modeling results and discusses the potential of each model. The final chapter summarizes the dissertation effort and highlights the research contributions.

The dissertation will present methods to process passively collected longitudinal GPS data, identify activity at the end of the trips, evaluate travel behavior variability, estimate the activity space and develop activity participation models. The developed methods will be applied to the data collected from the Commute Atlanta Study and the results will be presented. The dissertation expects to contribute to activity based modeling theory by showing the benefits of using travel behavior variability and activity space as variables that represent the household's variability-seeking nature and appetite for spatial extent. The travel variability and spatial extent are easy to estimate with data from most travel surveys and may help towards better travel behavior models.

CHAPTER 2

LITERATURE REVIEW

This chapter presents a literature review summary for topics relevant to this dissertation. The first section explores the literature on travel data collection methods and, in particular, cross-sectional and longitudinal travel data. The chapter then reviews the existing literature on travel variability studies. The third major section reviews the literature on activity purpose identification. This is followed by the section reviewing literature on activity space. The next section discusses the literature on activity participation modeling and the final section provides the summary of this chapter.

Travel Data Collection

The most common method for collecting travel data for use in transportation modeling and planning is to undertake one-day or two-day travel diary surveys. These travel diary surveys provide cross-sectional data for estimating average travel behavior within transportation analysis zones. The other primary data collection method is to collect travel data from smaller sets of households over longer time periods. The duration of these longitudinal travel surveys can vary from a few weeks to a few years. The longitudinal travel data are useful in assessing travel behavior at the disaggregate level and changes in travel behavior over time.

Cross Sectional Travel Data

The analysis of cross-sectional travel data assumes that the individual's day-to-day travel is fairly stereotyped, or habitual [13]. Utility maximization theory, i.e. individuals try to perform activities as efficiently as possible, forms the basis of data analysis [14]. In general, individuals try to satisfy their activity needs rather than optimize activities, and routine behavior is a stress-minimizing, satisfying strategy by eliminating the need for constant decision-making. This leads to the assumption that most people establish habitual behavior patterns. However, many activities of an individual occur in cycles that may be repeated daily, weekly, monthly, or even annually. Cross sectional travel data can present significant problems in drawing inferences from the data due to the temporal variability in travel behavior where potential errors may be associated with the temporal cycle over which an activity occurs. As noted above, typical travel surveys are only conducted in one-day or two-day durations. Hence,

- cross sectional travel data have analytical limitations with respect to behavioral dynamics in travel behavior [1]:
- Cross sectional data are inadequate to evaluate the response lags and response leads of behavioral adjustments to an event.
- Cross sectional data cannot capture habit persistence, where people exhibit routine behavior even after such behavior is no longer optimal (e.g. habitually shopping at one location even after a closer location with equal utility has opened).
- Cross sectional data are not useful in evaluating threshold or cumulative effects, where the magnitude of change associated with an event needs to be greater than or less than a threshold value for behavioral change.
- Cross sectional data cannot evaluate behavioral asymmetry or hysteresis, where people make asymmetric adjustments in behavior in response to symmetrically opposite events.
- Cross sectional data do not capture multiple equilibria, where multiple states of behavior are possible for any set of conditions. For example, a study on elasticity between car ownership

and income found that the elasticity changes over time and different household types exhibited different patterns in changes to elasticity [15].

GPS based Travel Data Collection

With recent advances in technology and reduction in costs of GPS data loggers, travel surveys using GPS are becoming common. Travel surveys that use GPS data are of three types. The first travel survey imitates the traditional travel diary. The participant carries a personal GPS device or his/her vehicle is instrumented with a device. The participant is also provided with a handheld computer in which he enters the trip characteristics at the end of the trip [16]. The GPS data and the data entered by the participant together form the data. In this method, the spatio-temporal accuracy of the GPS device and the human elements from the participant make the dataset comprehensive. However, this study cannot be done for a long period due to survey fatigue for the participants.

The second method is to passively collect GPS data by installing a GPS device on the participant's vehicle [16]. This method is very useful for safety studies where the primary interest is the vehicle parameters. Since there is no human input, identifying the activity type, driver, and passengers are difficult. Using this method, data can be collected for very long since the participants have no responsibilities. The Commute Atlanta Study is a good example of this travel survey [17]. The third method is a hybrid of the passive data collection with interim the travel diary surveys [16]. In this method, the participants will be shown their travel traces to help them recollect their trips. The information about the activity types at different trip ends from the travel diary will help in identifying the activity types on other days at the same location.

Longitudinal Travel Data

Longitudinal data are important from policy as well as analytical viewpoints. Jones and Clarke offer an excellent discussion on the significance and measurement of variability in travel behavior [2]. No matter how big the sample of cross sectional data, it cannot address variations in travel behavior over time. Longitudinal data facilitate the identification of cause and effect relationships that may be involved in behavioral dynamics [1]. Longitudinal data, typically collected in panel surveys, are efficient to collect with respect to survey cost and parameter efficiency. The Puget Sound transportation panel survey, which consisted of four waves of travel survey from 1989 to 1993, is a good example of longitudinal data collection using panel surveys. Longitudinal data collected using passive technologies, such as instrumented vehicles studies, can collect data for long periods of time without loss in accuracy or without yielding significant participant fatigue.

Longitudinal data and data collection processes have limitations. Passively-collecting longitudinal data requires state-of-the-art technology and highly-skilled labor. Longitudinal data collection in both panel surveys and passive instrumented surveys face continuity issues as subjects leave the study over time. The turnover of equipment during the course of a data collection period affects data collection efficiency. Data collection is also dependent on external services such as wireless and GPS services that may adversely affect the study when service interruptions occur. The cost per household for collecting longitudinal data is larger than cross-sectional data collection. When passive data collection technologies are used, it is more difficult to identify trip purpose when a trip ends at a new location. Standard CATI surveys provide the advantage of respondents stating the purpose of each trip in a completed survey. Inferring trip purpose from observed travel patterns will be discussed in detail in the next section. Finally, passive longitudinal data collection may not capture trips made by modes other than auto, unless cell phone or personal tracking mechanisms are employed, which can be a significant issue in behavior analysis.

However, with advances in computing resources and decrease in data transfer costs longitudinal travel studies using GPS are becoming common. Greaves, et al., developed a GPS based travel data collection system to passively collect GPS data over ten weeks in Sydney, Australia [18] The Commute Atlanta study is another instrumented vehicle study that collected GPS based travel data over a three year period from 500 vehicles with very little participant burden [17].

Travel Variability Studies Using Longitudinal Travel Data

Pas and Koppelman examined the determinants of data-to-day variability in individual's urban travel behavior from data collected during the Reading Activity Survey between January and March 1973 [4]. This research developed and examined two general hypotheses regarding the determinants of intrapersonal variability in urban travel behavior. The first hypothesis was that individuals with fewer economic and role-related constraints have more intra-personal variability in their daily trip frequency. The second hypothesis tested was that individuals who fulfill personal and household needs and do not require daily participation in out-of-home activities have higher levels of intrapersonal variability in their daily trip frequency. The study tested and verified the two hypotheses to be significant.

Schlich and Axhausen used six-week travel diary data to evaluate habitual travel behavior [14]. The paper found that the day-to-day behavior is more variable if measured with trip-based methods instead of time budget methods. The study observed that travel behavior is neither totally repetitious nor totally variable. The study concluded that the travel period observed should be at least two weeks to measure variability. A follow-up study in which Schönfelder and researchers at the Georgia Institute of Technology assessed Commute Atlanta data using the same methods found that the variability in travel activity was even larger in Atlanta than had been observed in Borlänge, with more new locations visited each day in Atlanta than observer in the European cities, indicating that continuous studies might need to run as long as 20 days in Atlanta to assess these patterns [3].

Using the same six-week travel diary data as Schlich and Axhausen, Susilo and Kitamura analyzed the day-to-day variability in an individual's action space [19]. This study's results show that out-of-home activity orientation and commitment influence the extension of action space. For workers and students, they observed that the spread of activity locations and the distance to these locations was stable from day-to-day (perhaps as if a constrained range is in place, given their existing constraints). The study found that random factors have dominant influence on non-workers weekday action spaces and on all individual's weekend action spaces.

Pas and Sundar examined the day-to-day variability in urban travel using a three-day travel data set collected in Seattle, WA [20]. Intra-personal variability was measured in terms of daily trip generation rates i.e. the mean intra-personal variance in daily trip generation rates. The study found considerable day to day variability in terms of trip frequency, trip chaining, and daily travel time. The paper found that day-to-day variability in travel time was similar to that of trip frequency.

Stopher et al. studied the variability of travel behavior by day of the week and among individual persons using panel data from South Australia [21]. The authors calculated the cumulative mean and variance of travel behavior attributes, such as number of trips, travel time and distance. They found that the cumulative means and variances of travel time and number of trips stabilized after 4 to 6 days of travel survey. However, cumulative mean and variance are not very useful for anything other than trip generation, certainly not for assessing repetitive behaviors and patterns in the travel stream.

The day-to-day variation of individual trip scheduling and route decision for the evening commute was studied by Hatcher and Mahmassani [22]. The detailed 2 week travel diary data from commuters in Austin, TX was used in this study. The study presented models to relate observed route and departure time switching patterns to the commuters' characteristics, such as workplace conditions, socioeconomic attributes, and traffic system characteristics. The study observed high variability of the daily departure time from work, which may be attributed to the trip-scheduling flexibility associated with this trip.

Li, Guensler and Ogle analyzed the morning commute route choice patterns using GPS based vehicle activity data from the Commute Atlanta Study [23]. This study presented a binary logit model for the choice of single route versus multiple route for morning commute based on route characteristics and household's socioeconomic characteristics. The study found that there was a strong relationship between morning commute route choice (single versus multiple routes) and commuters work flexibility, socio-demographic characteristics, and commute route attributes.

Hanson and Huff tested the assumption that individual's day to day travel is habitual and that a one-day travel record constitutes sufficient data for model building [13]. The authors analyzed the number of stops (trip ends) and number of journeys (from home onto the transportation network and returning home) per day[13]. The paper also looks at contingency tables between activity, mode, distance and location for repetition of travel patterns. The study found that even though there is substantial repetition of travel there is a noticeable difference in travel patterns even on weekdays. The study recommended extensive analysis of systemic variability in daily patterns.

Pas examined the effects of intra-personal travel variability on travel demand model goodness-of-fit [24]. The research found that there is considerable effect on the apparent goodness-of-fit of person level trip generation models estimated with cross-sectional data. The study found that the intra-personal travel variability is a significant part of total variability and the relative magnitude of the intra-personal variability varied with the trip purpose and population group.

Kang and Scott investigated the day-to-day variability activity time-use patterns of household members in the Toronto travel activity survey [25]. The study, using descriptive statistics and a series of daily structural equation models, found evidence of day-to-day variability in activity time-use patterns. The research found that weekday time-use patterns are very different than weekend time-use patterns. Results from this study also suggest that there is a higher proportion of intra-personal variability and lower proportion of inter-personal variability for joint activities when compared with independent activities [25].

Identifying Trip Purpose from GPS Data

In 2000, Wolf et al. undertook a proof-of-concept study with 30 participants on the possibility of using data from GPS data-loggers to identify all parameters including trip purpose [26]. The study overlaid the trip ends on a geographic referenced land use database. The study had to predominantly process the trip end data manually due to the variations in the land use data sources and accuracy issues. The study also found that only 22% of the trip ends needed follow-up questions to identify the trip purpose (i.e. multiple potential purposes were possible). It should be noted that the study had only 30 test subjects and their trips were monitored for three days; hence, it was possible to manually assign trip purpose for trip ends that were not automatically assigned to a land use parcel. Manual assignment of trip purpose for travel surveys involving 10,000 households is simply not practical

Schönfelder et al. explored the potential of using automatic GPS for travel behavior analysis in 2002 using the data collected from Borlänge between 1999 and 2001[16]. The Schönfelder study processed raw GPS data to first identify trips and trip end locations. To identify trip purpose, the study used the underlying land use parcel data, survey information about occupation and habitual patterns to travel. The Schönfelder study noted that for different land use blocks, the radii to search for the trip ends are different depending on the neighboring land use pattern and parcel size.

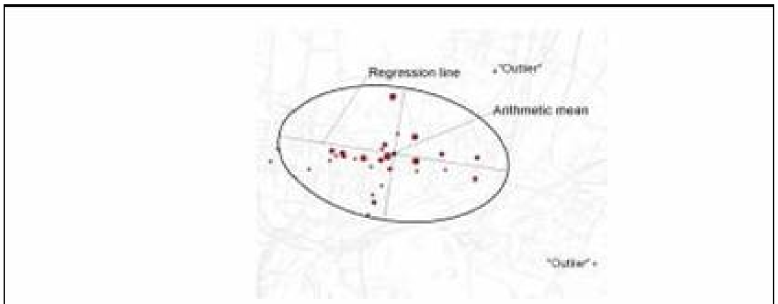
Lu, et al., explored the feasibility of automating trip purpose identification using machine learning algorithms with land use data [27]. The study compared three machine learning methods, viz. cluster based land use coding method, closest point of interest land use coding method, and the metalearner method, finding that the metalearner machine learning method performed best [27].

Activity Space

The activity space is the spatial area in which a household undertakes a large percentage, typically 95%, of its activities. The activity space for households can be estimated by the following methods.

Confidence Ellipse or Prediction Interval Ellipse

Confidence ellipses are analogous to the confidence interval of univariate distributions as the smallest possible (sub-)area in which the true value of the population should be found with a certain probability (e.g. 95%) [28]. The measure of activity space is the area within the confidence ellipse. The ellipse shows the dispersion of the locations visited by the household. ArcGIS 9.2 provides spatial tools to estimate the confidence ellipse [29]. The axes of the confidence ellipse is defined by estimating the standard distance along the x and y directions separately. The shape and orientation of the generated ellipse provide the spatial trend of the location data. Figure 2.1, shows a confidence ellipse with all the locations visited by the household.



Dots show location and intensity of observed activity locations of one respondent

Figure 2.1 Confidence Ellipse for Activity Space [28]

The Covariance matrix of all the locations visited by the household calculates the confidence ellipse [28].

$$S = \begin{pmatrix} S_{xx} & S_{xy} \\ S_{yx} & S_{yy} \end{pmatrix}$$

Where each covariance is defined as [28]

$$S_{xx} = \frac{1}{n-2} \sum_{i=1}^{n} (x_i - \bar{x})^2$$

$$S_{yy} = \frac{1}{n-2} \sum_{i=1}^{n} (y_i - \bar{y})^2$$

$$S_{xy} = S_{yx} = \frac{1}{n-2} \sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y})$$

The determinant of the Covariance matrix is [28]

$$|S| = S_{xx}S_{yy} - S_{xy}^2$$

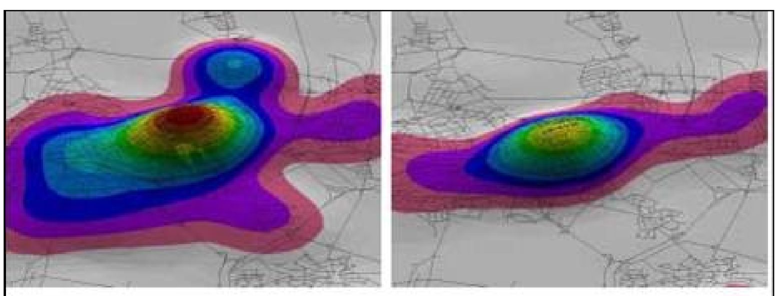
The area of the ellipse is given by [28]

$$A = 6\pi|S|^{\frac{1}{2}}$$

The ellipse tends to work well in representing travel space, given the presence of linear freeway structures (suburb to city and suburb to suburb) and linear major arterial structures in major US cities. Alternative spatial structures can be used to represent travel space, such as the non-parametric Kernel Density approach.

Kernel Density

Kernel densities are the transformation of a point pattern, such as a set of activity locations, into a continuous representation of density in a wider area [28]. The kernel density area is the sum of all areas with at least certain non-zero probability of activity occurrence. The estimation of kernel density is a smoothing technique, which generalizes point to the area in which they are found. Figure 2.2 shows the kernel density for the activity space of a household.



*Total Karlsruhe sub-sample; Left: leisure activities, right: work

Figure 2.2 Kernel Density for Activity Space [28]

There are multiple approaches to the estimation of kernel density and one method is the Fixed Kernel Method. In this method, a variably distributed kernel function is placed over each data point. The sum of the overlapping values provide the density estimate. Considering a grid structure in which single points are represented by grid cells (e.g., a raster representation in GIS), the base kernel density is [28]

$$\lambda(s) = \sum_{d_i} K\left(\frac{d_i}{\tau}\right)$$

Where

λ the density estimate at grid cells

τ the bandwidth or smoothing parameter

K kernel function which should be further specified

d_i distance between grid point s and the observation of the i^{th} event.

Kernel density is commonly used when exploring density of events such as crime locations, or locations visited etc. Kernel density methods allows spatial arithmetic operations between different layers and also can be used for direct comparison. Chapter 8 discusses the benefits and limitations of the Confidence Ellipse and Kernel Density methods in detail.

Minimum Spanning Trees

Minimum Spanning Trees assume that there may be correlation between the frequency of using a network link and the individual's knowledge of the surrounding area. The identified area of perception widens around home with several further visited activity locations and narrows along the links to the other (main) activity centers [28]. The measure of activity space is the weighted overall length of used network by the household. The underlying assumption in this method is that activity occurs in areas that are known to the individual and the more an individual travels along a route the more they are exposed to activities along the route. Figure 2.3 shows the minimum spanning network for a typical household.

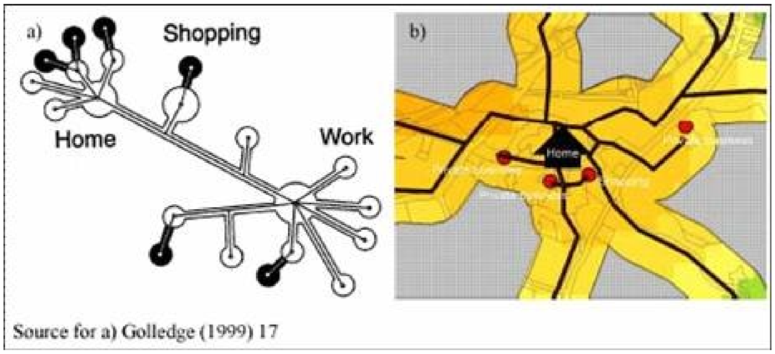


Figure 2.3 Minimum Spanning Trees [28]

Summary of Activity Space Methods

The three primary activity space estimation methods discussed in this section are summarized in Table 2.1. The confidence ellipse and kernel density methods provide area estimates while the minimum spanning tree provides a linear distance measure. The confidence ellipse can cover vast spaces where no activity occurs and may overestimate the activity space. Kernel density on the other hand, does not take into account the spatial dispersion of activities and can underestimate the activity space of an individual. The minimum spanning tree does not take into account the buffer around an individual which influences his or her activities. This buffer distance can vary depending on whether the individual is driving or participating in another activity. An activity space that estimates the true extent of an individual's activity can provide the spatial appetite of that individual and can be used in activity participation modeling to better explain the travel behavior.

Table 2.1 Activity Space Estimation Methods

Measure	Indicator
Confidence Ellipses	Area of Ellipse
Kernel Density	Share of covered area in reference area, from cells that exceed specified threshold
Minimum Spanning Tree	Weighted overall length of used network

Activity Participation Modeling

Activity based approaches to travel demand modeling are based on the view that travel is derived from the need of the households to participate in activities that are spatially distributed [5, 30, 31]. Traditionally discrete choice modeling and hazard duration models have been used to model travel behavior [5, 32].

The multinomial logit method is the widely used discrete choice modeling structure in travel behavior analysis [5]. The discrete choice models follow the utility maximization conceptual framework which is based on the economic theory of choice that individuals make choices to maximize utility for themselves [30]. This method assumes that within the time and cost constraints imposed by an individual's budget, they chose to spend time in activities that maximizes their goals and desires. The utilities are assumed to be dependent on the duration, time of day and frequency of the activities to explain the change in activities over time [30]. The validity of the utility maximization assumption has been questioned since individuals decision-making process relies more on habitual patterns than utility maximization and it is recommended to at least modify this assumption [30, 33, 34].

Hazard duration models focus on the end of duration of an activity given that a certain time has already elapsed in that activity [5]. The concept of conditional probability of failure or termination of an event recognizes the influence of time already spent on that activity. Proportional hazard models are commonly used to examine covariates that influence duration time. This method assumes that covariates act multiplicatively on some underlying hazard. The two main issues for using this method is selecting the appropriate distribution of the duration and the assumption about unobserved heterogeneity or difference in duration across individuals. The distribution of the hazard may be assumed to be parametric or non-parametric. The choice of parametric distribution of the hazard can lead to inconsistent baseline hazard estimation and covariates effects if the assumptions of that parametric distribution are not met [5, 35]. The failure to control for unobserved heterogeneity can lead to severe bias in the duration dependence and the estimates of the covariates effects [5].

Structural equation modeling has been used more recently in activity participation modeling. The direct relationship between activity demand and the need to travel to reach the activity site, inter-relationships between the different activities and impacts of travel time on activity time can be modeled using the structural equation modeling methods [36]. Golob states the considerable potential of structural equation modeling in activity based travel demand modeling to create a comprehensive framework that captures direct relationships between activity demand and the need to travel, interrelationships between the need to participate in different activities and the feedbacks from travel time to activity time [36].

Structural Equation Model

Structural equations modeling is the union of latent variable or factor analytic approaches of psychology and sociology, and the simultaneous equation methods of econometrics [36]. Structural equations modeling has been applied by social and behavioral scientists to study casual relationships [37]. The structural equation modeling is used to capture the casual or regression effects of exogenous or independent variables on endogenous or dependent variables and the casual influences of endogenous variables on each other. Structural equations model has advantages when compared to other linear-in-parameter statistical methods. It can treat both endogenous and exogenous variables as random variables with errors of measurement, have latent variables with multiple indicators, separation of measurement errors from specification errors, test model overall instead of individual coefficients, model mediating variables, model error term relationships, test coefficients across multiple groups in a sample, model dynamic phenomena such as habit and inertia, account for missing data and handle non-normal data [36].

Structural equation model (SEM) is a set of simultaneous equations specified by direct links between variables [38]. SEM account for modeling of interactions, nonlinearities, correlated independents, measurement error, correlated error terms, multiple latent independents each measured by multiple indicators and one or more latent dependents also each with multiple indicators [39]. There are three parts to SEM: a measurement sub-model for the endogenous variables; a measurement sub-model for the exogenous variables and; a structural sub-model involving latent variables. A SEM without latent variables is defined by

$$y = By + \Gamma x + \zeta$$

Where

y column vector of p endogenous variables,

x column vector of q exogenous variables, and

ζ column vector of the error terms

B matrix (p*p) of direct effects between endogenous variables

Γ matrix(p*q) of regression effects of the exogenous variables

Let

Φ Covariance matrix of x

Ψ Covariance matrix of ζ

Θ population covariance matrix of observed variables expressed in terms of unknown parameters in B, Γ, Φ, and Ψ matrices.

The parameter Θ can be estimated by minimizing the discrepancies between the sample covariance matrix (S) and the population covariance matrix $\Sigma(\theta)$. The fitting function for structural equations maximum likelihood estimation is

$$F_{ML} = log\left|\sum(\theta)\right| - \ log|S| + tr\left[S\sum(\theta)\right] - (p+q)$$

With the assumption of multivariate normality, (n-1) F_{ML} is χ^2 distributed. This provides a test of model rejection and criteria for testing hierarchical models. The maximum likelihood estimation method is commonly used in travel behavior research and has been found to be robust against violations of multivariate normality [36]. The details of the activity participation model development using structural equations model will be discussed in detail in Chapter 10.

Travel Behavior using Structural Equations Modeling

Golob and McNally used a structural model to assess activity interactions between heads of households and thereby explain the household demand for travel [40]. The authors used data from the 1994 Portland Activity and Travel Survey. The results suggest that a feedback mechanism need to be introduced in trip generation models to reflect the effect of activity frequency and duration on the generated travel. A household that has long commute times will need to have compensatory reduction in and travel to other types of activities.

Golob used structural equation methods to model household mobility decisions as derivative of activity participation decisions [41]. The study used activity participation and activity locations as endogenous variables, where the activity location is dependent on the activity participation selection. The study used the two day activity diary data for male and female head of households and associated accessibility that was collected in Portland, Oregon in 1994. The developed model had the potential to forecast the effects of accessibility, in-home work on travel demand, car ownership and vehicle miles traveled. The developed model was able to forecast the effects of activity participation and mobility factors that are difficult to include in trip based models [41]. The structural equation model that was developed can test the effects of additional exogenous variables such as accessibility measure on activity and mobility demand, expand the set of mobility variables to explore how trip generation is related to activity participation, and also divide activity demand by weekday and weekend.

Lu and Pas developed, estimated and interpreted a structural equations model that relates socio-demographics, activity participation and travel behavior [42]. The study found that travel behavior may be better explained by using activity participation as endogenous variables, rather than using socio-demographic variables alone. The research also shows that examining the direct and indirect effects in the model system help better capture and understand the relationships among socio-demographic variables, activity participation and travel behavior.

Kuppam and Pendyala did an exploratory analysis of commuter's activity and travel patterns to investigate and estimate the relationships between demographics, activity participation and travel behavior using structural equations modeling [37]. The study used data from the activity based travel survey data collected in the Washington DC metropolitan area. The authors found strong relationships between commuter's demographics, activity engagement information and travel behavior.

A series of structural equation models that capture relationships among socio-demographics, activity participation, and travel behavior for each day in a week for developing country are presented by Chung et al in 2002 [38]. From the empirical results, Chung et al conclude that there may be similar relationships between socio-demographics and travel behavior in developing and developed nations [38]. The study confirmed that weekend and weekday activity patterns are significantly different. Finally, the publication presents and explains the relationships among socio-demographics, activity participation and travel behavior from the direct, indirect and total effects in structural equation model.

A multilevel Structural Equation Model was used by Chung et al in 2004 to handle the hierarchical nature of the household and individual data [43]. The authors use this modeling approach to try to better explain the complex relationships among socioeconomic factors for individuals and households, activity participation, and travel behavior using data from the third wave of the Puget Sound Transportation Panel. Empirical results from the study found high interdependency for leisure activity duration and total travel time duration for household members within a household. The results suggest that multilevel structural equation modeling approach to be an appropriate modeling methodology in capturing the relationship between activity duration and the total travel time [43].

Kang, et al. developed a structural equations model of daily time allocation to analyze activity patterns of household heads [44]. The model explored intra-household interactions by differentiating joint activities and independent activities within the model. The study used the data from 2002-2003 Toronto Travel-Activity Pane Survey (TTAPS). The study found significant trade-offs between independent and joint out-of-home activities. The presence of children and higher car ownership were found to have negative impact on joint activities [44]. The ability of structural equation models to incorporate the joint and independent activities and model the interactions between the two variables provides a strong framework to explore casual relationships that hazard duration models and multinomial logit models do not provide.

Weis and Axhausen studied the aggregate effects of changed generalized costs of travel on traffic generation, the propensity for out-of-home activities on a given day, number of trips, and the resulting total out-of-home duration and distance traveled using structural equations model [45]. They used a pseudo panel data set constructed using the Swiss National Travel surveys conducted since 1974. The study found that the generalized costs and accessibility elasticities were substantial even after correcting for socio-demographic effects. Structural equations modeling provided the framework to integrate generalized cots of travel and accessibility as variables with activity demand models and study their elasticity.

Wang and Lin studied the impact of the built environment and the socio-demographic characteristics on travel behavior using structural equations model using data from travel diary survey conducted in Hong Kong [46]. People living in private and public housing were found to have different[47] built and social environments in Hong Kong. The study found that the built environment significantly determines social environments which in turn influences travel behavior.

Abreu e Silva, et al., studied the relationship between travel behavior and land use patterns using structural equation modeling framework [47]. The research effort found that people with different socio-economic characteristics tend to work and live in places characterized substantially different urban environments. The effects of land use influence variables describing long term decisions like commuting distance, and car ownership [47].

Summary

This chapter reviewed the existing literature to better understand travel survey data collection and analysis methods that are relevant to this dissertation. The first section explored the travel diary systems used to collect travel data. The differences between cross-sectional travel data and longitudinal data are reviewed. The benefits and limitations of using longitudinal data were also discussed. Existing literature on travel behavior variability studies were presented in the next section. The literature on activity purpose identification methods were then presented. This is followed by the section reviewing literature on activity spaces and the methods to estimate them. The last section discussed the existing literature on activity participation modeling methods and previous literature on modeling results. The section also briefly reviewed the structural equation modeling methods that will be used in this dissertation. The methodologies will be further explored in Chapter 10.

The literature review has identified existing literature on the methods to identify trip purpose which is a very important step in making the Commute Atlanta dataset useful for modeling. This chapter also explored varies methods used by researchers to quantify travel behavior variability and its relationships with socio-demographic variables which will help the dissertation in identifying the right methodological approach to use. The methodologies to estimate activity space and their applications have been explored to help in developing the right activity space measure. Finally, activity participation modeling efforts have been reviewed and structural equation modeling method has been explored in detail. Studies using structural equations modeling and their ability to infer complex inter-relationships and casual relationships between variables have been examined to better understand the application of this method to build activity participation models. Structural equations modeling are more intuitive, can create a framework that captures direct relationships between activity demand and the need to travel, interrelationships between the need to participate in different activities and the feedbacks from travel time to activity time. Structural equations model also provide the ability to integrate trip generation as part of the activity participation modeling process in an elegant way.

CHAPTER 3

DATA COLLECTION

This chapter presents the data employed in this dissertation research. The analyses primarily use longitudinal GPS data collected as part of the Commute Atlanta study [17]. The GPS data collected by University of Minnesota as part of their I-35W travel behavior study are used in a case study to evaluate the trip purpose methodology developed in this dissertation [12]. The Commute Atlanta data collection, processing, filtering, and trip chain identification methodologies are discussed first. Next, the chapter describes the data collection by the University of Minnesota study and the processing techniques developed to make the data usable.

Commute Atlanta Data

The Federal Highway Administration's Value Pricing Program and the Georgia Department of Transportation funded the Commute Atlanta Study. The primary objective of this study was to assess the effects of converting fixed automotive costs into variable driving costs[17]. The research objective was to test the hypothesis that given a per mile pricing system, participants will modify their driving patterns in an effort to reduce their total costs, pocketing the savings.

The Commute Atlanta study collected data from approximately 500 vehicles present in 268 households in the metro Atlanta area. Figure 3.1 shows the spatial distribution of the participating household locations. A random stratified sampling technique was used by the study to recruit participants by income, household size, and vehicle ownership to counter low recruitment in certain strata that can occur in travel diary studies, especially of lower income households [17]. However, even with a second round of targeted sampling, due to the instrumented nature of the study, the participation of lower-income households was lower than targeted and participation of households with income greater than $100,000 was higher than targeted.

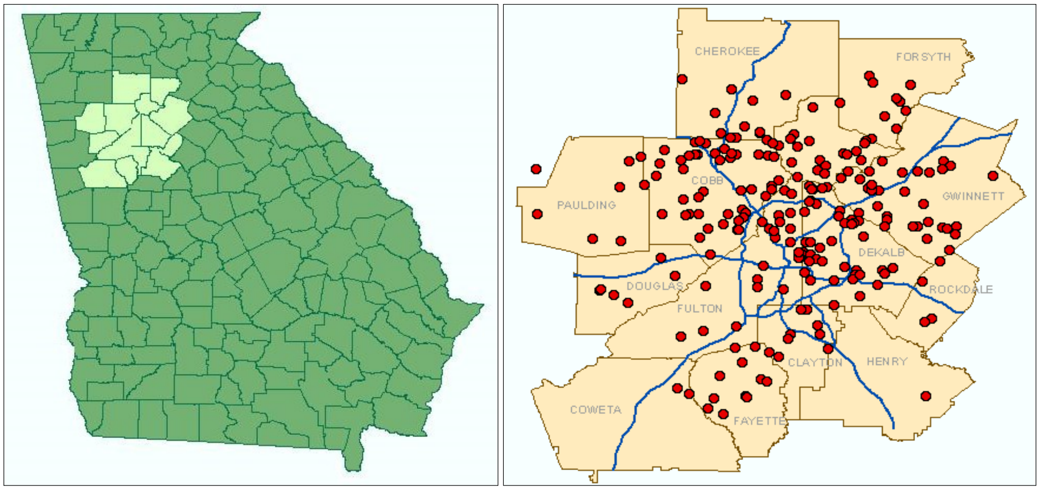


Figure 3.1 Commute Atlanta Participant Household Locations

Instrumented vehicle data were collected from August 2003 through June 2006. The installation of the Georgia Tech trip data collectors started in the fall of 2003. In the first phase the baseline travel data were collected for households from fall 2004 through summer 2005. The second phase implemented cent-per-mile incentives from October 2005 through June 2006.

One major conclusion of the study was that the variability in household demographics over time significantly affected the intra-household travel behavior variability [48]. About 70 percent of the households experienced demographic and vehicle ownership changes in the baseline and pricing periods. With the large variability in travel behavior, it proved impossible to determine definitively any household response to the pricing incentives. The study concluded that much larger sample sizes and improved survey design are required in longitudinal studies to measure the effects of pricing on travel behavior [11].

Data Collection

The data collection in the Commute Atlanta study had two elements viz, the passive GPS data from the instrumented vehicle fleet and the household socio-demographic data from surveys.

Georgia Tech Trip Data Collector

The Georgia Tech trip data collector (GT-TDC) shown in Figure 3.2, was designed in the Drive lab to collect GPS data, wheel-tick speed data, and data from the on-board computer. The GT-TDC embedded a 386 Linux Central Processing Unit (CPU) with 8 megabytes of Read Access Memory (RAM). A cellular trans-receiver to transfer the data remotely, SiRFstar II GPS chipset, on-board diagnostics engine computer connection, and a vehicle speed sensor were also part of the GT-TDC.

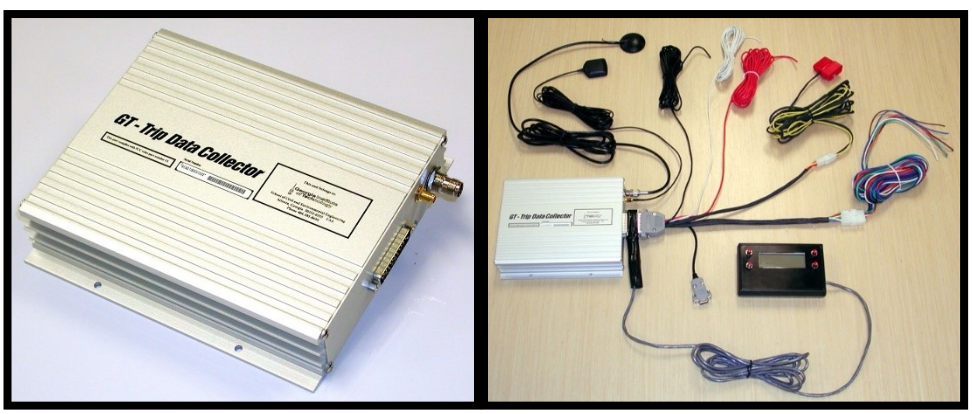


Figure 3.2 Georgia Tech Trip Data Collector (GT-TDC) and wiring harness

Figure 3.3 shows the data collection schema of the Commute Atlanta Study. The GT-TDC was installed on vehicles that drive more than 3000 miles annually in participating households. The units were installed usually under the front seats to protect them during crashes and other hazards, and to not interfere with driving. The GPS antenna was installed on the windshield facing the open sky to get good GPS signals. The GT-TDC was connected to the vehicle ignition line and is powered when the vehicle is turned on. The GT-TDC records second-by-second the GPS data, on-board vehicle diagnostics data, and the wheel-tick data and stores it into the trip files. At the end of the trip, when the ignition key is switched-off, the trip file is closed and saved in a flash drive on the device. The trip files were automatically transferred every night back to the research lab using file transfer protocols (ftp) over wireless communications. The trip files are processed and stored in the secure server. A trip file contains data between engine key-on and key-off events.

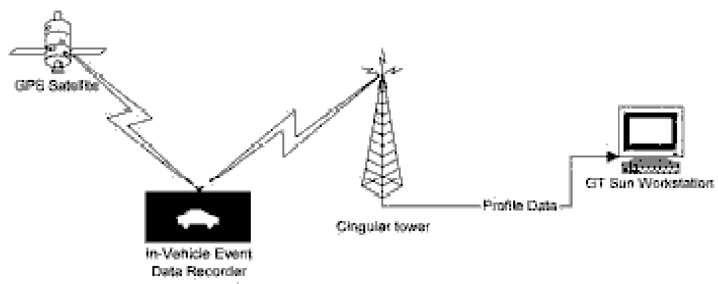


Figure 3.3 Commute Atlanta Data Collection Schema

Household Socio-Demographic Data Collection

The household socio-demographic data were collected using surveys. At the time of recruitment, respondents were surveyed on their socio-demographic data and efforts were made to recruit 35 to 40 participants in each recruitment strata defined by the household income, size and vehicle ownership [17]. In addition to the survey at the time of recruitment, the study requested the participants to complete two-day travel diaries in the summer of 2003 and spring of 2004. As part of these surveys, the socio-demographic data were collected and the changes were updated.

Due to the longitudinal nature of the data collection, most households had changes to their socio-demographic characteristics. It was observed that 70 percent of the 95 households that participated in the second phase had changes to the socio-demographic characteristics during the baseline and pricing period [48]. Because this was the first study with data collection over three years, the study was not initially designed to capture these changes with regular follow up surveys. The research team realized this issue at the beginning of the pricing period and mailed out socio-demographic information to each household every quarter for verification.

The research team updated changes to vehicle ownership more accurately than the other household variables since the households called the research team to uninstall the GT-TDC from the sold vehicles and install in newer vehicles. However, the households did not automatically update other changes in work place, work status, student status, household income, household size, etc. The study captured most of the changes in the socio-demographic status with the verification surveys for the 95 households that participated in phase II.

Data Processing

The first step in data processing was the quality check on the data uploaded by the GT-TDC to the secure server. Automated scripts checked the compressed binary files for quality using an MD5 (message-digest algorithm) check sum which is a small size datum from an arbitrary block of data to detect errors during transmission. The binary files are then decompressed to extract the raw data. The raw data were stored in American Standard Code for Information Interchange (ASCII) files and a compressed archive was used for backup.

The next step was to process the data in a Geographic Information Systems (GIS) environment and match the GPS records to the actual routes taken by the vehicle. The GPS records at the beginning of a trip are usually invalid due to the cold start of the GPS chipset [49]. Therefore, to identify the correct origin of a trip, the last GPS record of the previous trip is added to the beginning of each trip[23].

Then the trip files are converted into point features in the GIS environment and using a buffer zone assigned to the nearest link on the road network. The route processed data need to be corrected to fix errors when they are assigned to neighboring road network links or intersection where GPS data records could be assigned to cross-streets. The data are smoothed by examining the link that was travelled before and after a stop and the right network link is assigned for those GPS records.

The socio-demographic data collected from the participants were stored in a relational database that was custom-designed for this study. The relational database had the ability to track changes in socio-demographic variables over time by using temporal variables as part of the tables. The relational database also had the information on the instruments, their status, and which vehicle they were installed. The results of quality control quality assurance (QA/QC) methods on the dataset identified the periods the GT-TDC units were not working and this information was also stored in the relational database. The GT-TDC used to collect data would occasionally break down and needed replacement. The travel data during the time between identification of the problems with the GT-TDC and replacement of those units were lost. These issues were identified actively during the baseline period and only households that had complete travel data in the baseline period participated in the pricing part of the study. The households that participated in the pricing period where asked to verify their socio-demographic data every quarter and that information was recorded in the relational database.

Attributes

The attributes that were collected in the Trip data files included the data type, date, time, latitude, longitude, speed, heading, number of satellites, positional dilution of precision (PDOP), and the on-board diagnostic computer data if available. After route processing, the road-network link identity, closest mile marker, annual average daily traffic (AADT), roadway classification, speed limit, and truck are associated with the second by second data.

The socio-demographic attributes that were collected by the surveys included person, vehicle, and household characteristics. Household characteristics included address, contact information, income, number of people, and number of vehicles in the household. Vehicle characteristics included make, model, model year, body type, fuel type, and whether the vehicle was used for commercial purposes. For each person in the household, their age at the time of recruitment, birth year, gender, license status, student status, work status, ethnicity, school locations, work locations, highest education level attained, and percentage time they drive each vehicle in the household.

The number of children, number of workers, number of students, income group, age group, number of vehicles instrumented, vehicle model year group, and Georgia Tech Sample number were derived from the other socio-demographic variables. The trip characteristics such as start date, start time, end date, end time, distance, duration, whether the trip occurred inside the Metro Atlanta, soak time or activity time between trips, and the day of week were derived from the trip information collected by the GT-TDC.

Research Dataset

The primary objective of this dissertation is to use longitudinal variability of spatial activity and trip characteristics in activity participation modeling as surrogates of variability-seeking nature and spatial appetite of the participants. Activity participation modeling requires the complete travel data for the household members and changes in the demographic characteristics.

This dissertation uses the baseline, and pricing data from households that participated in the pricing phase of the Commute Atlanta data because the travel data are complete and socio-demographic changes verified.

The dissertation is proposing to use the intra-household variability in travel behavior as surrogate measures of the variability-seeking-nature of the household and spatial appetite of the households that are not captured by travel diaries. The phase II dataset under consideration has two sets of nine months data (baseline and pricing), totaling eighteen months of processed data. However, if all data for a household are considered together, the changes in socio-demographic variables will affect the travel variability and the variability attributes will not be valid surrogates. Schlich and Shöenfelder found that a minimum period of two weeks of travel data observations are required for studying intra-personal travel variability and recommend longer periods to effectively study the changes in weekends [3, 14]. Therefore, the dissertation will examine the travel behavior of each household by each calendar month to ensure that the period under consideration is long enough to study intra-household variability and so as not be adversely affected by the changes in socio-demographic attributes.

To make the dataset useful for analysis, the trip data and the socio-demographic variables must be linked. By running a query on the relational database based on the date of each trip, the socio-demographic attributes for the primary driver, vehicle and household are obtained and attached with the trip attributes. The trips table generated is aggregated by day and by vehicle to generate the vehicle-day summary table, which is again aggregated each day and by household to generate the household-day summary table. These three tables together constitute the primary dataset for this dissertation.

Data Filtering

The data from the Commute Atlanta baseline and pricing periods for the 95 households that participated in phase II of the Commute Atlanta study were used in this research. Data had to be cleaned to remove bad GPS data and vehicle engine starts without movement (non-trips). The quality of some of the GPS data were affected by cold-start where the device does not have information on it last location, visible satellites, and almanac information of satellites. Invalid GPS data were recorded until the GPS receiver locates the satellites and locks the vehicle position. In some cases, due to broken GPS receivers or GPS antennas, or very short duration trips during which no satellite lock is obtained, the devices record complete trips without any valid GPS points. Trips without valid GPS records do not provide accurate temporal and spatial data. Moreover, when there are significant number of trips in a household month that do not have any valid GPS records, the spatial variability estimation will be incomplete. The dissertation assumes that at a minimum 90% of the trips in a household-month need to have valid GPS records to be included in the analysis. Out of 1530 household-months collected, 171 household-months included more than 10% of the trips with invalid records. These 171 household-months (11.1% of household-months collected) were filtered from the final dataset. Another 13 household-months were found to have inconsistent dates entered into the socio-demographic tables and were filtered from the dataset.

The GT-TDC recorded all engine starts; however, some engine starts do not result in any trips. These trips need to be filtered from the dataset as they do not change the location or the purpose of the activity. The distance were calculated by Kalman filtering method for the Commute Atlanta dataset to smooth the invalid GPS records, random GPS errors and GPS wandering when the vehicles are stationary [49]. All trips that have less than 0.1 miles were assumed to be engine-starts without any trips and those trips were filtered from the dataset. This will eliminate some very short trips (e.g. relocation of a vehicle in a large parking lot, or visiting a neighbor down the street); however, these trips are not expected to be significant within the context regional travel demand analysis. The vehicle-day trips summary and household-day trips summary were recalculated to reflect the engine starts that were filtered.

Trip Chain Identification

A trip file corresponds to engine on and off for a vehicle. When a vehicle makes multiple trips without stopping the engine (trip chaining), only one trip file is generated. Trip chains account for activities such as pick-up, drop-off, drive-through ATM, post-box, and drive-through restaurants. These activities have very small dwell times and typically may require the driver to alter the routes to accommodate these activities. Therefore, it is important to also capture the activities that occur without an engine-stop.

To identify trip chains in passively collected data, researchers split the trips into trip components based upon non-movement of the device/vehicle. Previous researchers have used dwell times ranging from 30 to 120 seconds to identify trip chains [50-52]. The use of short dwell times may capture stopping at traffic lights or in congestion as a trip chain. Using longer dwell times, such as 120 seconds, may miss activities that require less time, such as pick-up and drop-off. Pearson, after studying the data from the 1997 Austin Household survey and other studies, concluded 45 seconds was suitable for dwell time [53]. Wolf found that a dwell time of 120 seconds for Atlanta was appropriate based on signal cycles and stops in congestion [52].

The GPS data in the Commute Atlanta study are processed to a route network [23]. This information can be leveraged to help identify traffic light stops and stops in congestion and limit these events from being from being identified as activity locations. The route processed GPS second-by-second records are grouped into chunks of on-road-network and off-road-network data. For each group of data records, the number of consecutive GPS records that have speeds less than 5mph are determined. A limiting speed of 5mph was used to account for GPS wandering and it is unlikely that a vehicle would traverse a reasonable amount of distance at speeds lower than 5mph. If the number of consecutive records with speeds less than 5mph is greater than dwell time then the location was concluded to be an activity location. This process was repeated for all the groups of off road-network data within the trip and all trip chains are identified. The limitation of this algorithm is that it does not have the ability to identify pick-up, and drop-off activities that happen on the road-network. Another limitation of this algorithm is that it could identify someone pulling over to read maps or answer a phone call as stopping for an activity. The dissertation developed an algorithm shown in Figure 3.4.

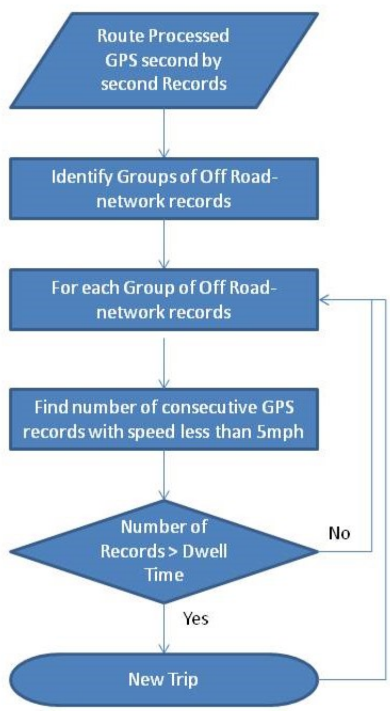


Figure 3.4 Algorithm to Identify Trip Chains

Table 3.1 shows the increase in the number of trips when dwell times of 30, 60, 90 and 120 seconds were used. Pearson found that 45 seconds was appropriate for Austin, TX while Wolf found that 120 seconds was more appropriate for Atlanta, GA due to signal cycles and congestion levels[52, 53]. However, Wolf did not use the underlying road-network to identify the activity locations. This dissertation uses the dwell time of 30 seconds because the Commute Atlanta data was processed to road networks, and pick-up and drop-off activities tend to be short. Using the shorter dwell time helps capture all trip chains from the Commute Atlanta data.

Table 3.1 Number of Trips with Different Dwell Times

Case	Number of Trips	Percent Increase
No Trip Chaining	243533	0.00%
Dwell Time = 120 secs	249676	2.52%
Dwell Time = 90 secs	253741	4.19%
Dwell Time = 60 secs	261542	7.39%
Dwell Time = 30 secs	282166	15.86%

University of Minnesota I-35 Bridge Study

The University of Minnesota conducted a travel behavior study on the use of the I-35W Bridge, which reopened in September 2008 after its fatal collapse in 2007. The study included 46 participants who commute across the I-35W Bridge. Each participant's vehicle was outfitted with a vehicle-based GPS system provided by Vehicle Monitoring Technologies, Inc. that transmitted second-by-second vehicle position data to a central server in real-time using GPRS/GSM communications. Data were collected from September 2008 through December 2008. The GPS device also transmitted engine on/off reports to the server[12].

The demographic data of the entire household were not collected during recruitment and only the individual participant's data are available. Data of other drivers in the household, work location of other family members and school locations of children are not available in this dataset.

The raw GPS data were processed to trips and maps containing trip traces were automatically created by the server. The participants could log in to a website to see their travel journal and complete an online travel diary to provide trip purpose and other trip-related details. The travel survey was a hybrid of passive data collection with interim requests for travel diaries. Each participant was requested to fill 6-14 days of travel diary through the study period.

Figure 3.4 shows the screen snapshot of Trip purpose recording page. One participant did not receive the travel diary requests because of an e-mail address error. The rest of the participants completed 94% of the travel diary requests. Some of the participants were apparently intrigued enough by the new travel diary system that they voluntarily completed travel diaries for additional days, without being asked to do so. This led to an unexpected data provision rate of 200% [12]. That is, participants reported trip purpose details for twice as many trips as they were asked to provide data for. Participants recorded the trip purpose data for more than 4,300 trips. This study employs only the 2185 trips for which trip purpose was requested.

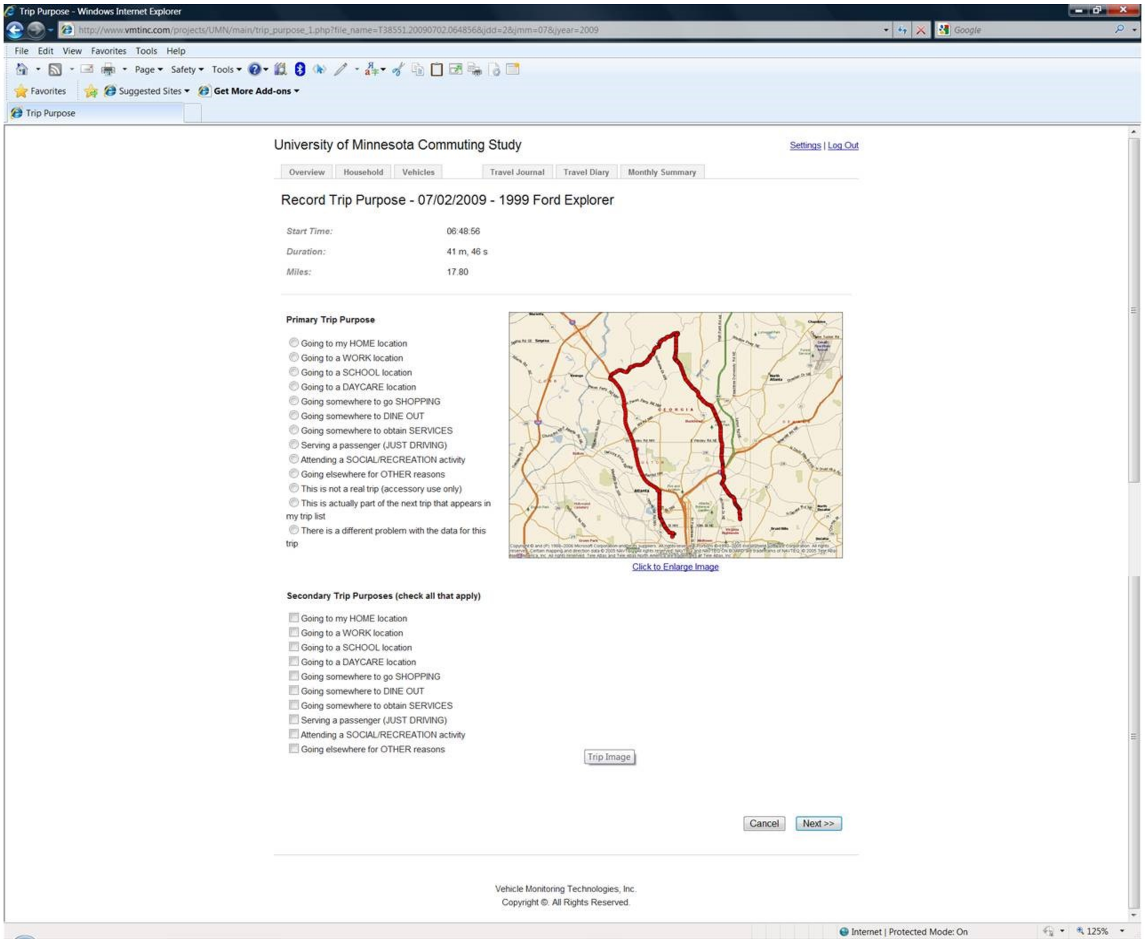


Figure 3.5 Screenshot of the Primary and Secondary Trip Purpose Recording Page

Summary

This chapter presented the data collected from Commute Atlanta study that will be used in this dissertation. The data collection instrument, the Georgia Tech trip data collector (GT-TDC), was designed in the Drive lab to collect GPS data, wheel-tick speed data, and data from the on-board computer. The household socio-demographic data were collected using surveys. Sophisticated data processing algorithms were used to process the raw GPS data into useful trip data. The trip data were then joined with the household demographics that were stored in a relational database.

The research dataset includes the baseline and pricing data of the 95 households for which complete travel data and socio demographic data were available. The data were cleaned to filter engine starts (non-trips), and households with more than 10% of bad GPS data. Households with inconsistent socio-demographic data were also eliminated. Various trip chain identification methods were explored. A methodology that took advantage of the route processing done to the Commute Atlanta data was developed and a threshold of 30 seconds of stopping while the vehicle was out of the road network was used to identify trip chains.
Data collected by University of Minnesota during a travel behavior study on the use of the I-35W Bridge, which reopened in September 2008 after its fatal collapse in 2007 were used to evaluate trip purpose identification methodology. Participants were apparently intrigued enough by the new travel diary system that they voluntarily completed travel diaries for additional days. This led to an unexpected data provision rate of 200%. The dissertation will use the 2185 trips from that study for which travel diary were requested.

CHAPTER 4

ACTIVITY TYPE IDENTIFICATION - METHODOLOGY

The methods described in Chapter 3 were applied to the Commute Atlanta data to generate trips with information such as origin, destination, departure and arrival. Passively collected GPS travel data are spatially and temporally accurate and can be collected over very long periods. Spatial data at this frequency and quality cannot be accomplished using traditional travel diaries. However, passive data collection does not include interaction with participants; hence, some of the information related to travel such as the activity at the end of the travel are not readily available. Accurate spatial and temporal data from passively collected GPS data are not useful in modeling unless the activity at the end of the trip can be deduced in other ways. Therefore most activity-based models developed by and for transportation planning organizations do not directly incorporate GPS data, even though GPS data has the accuracy and higher resolution that is required for those models[11]. Most model developers are using GPS data streams to supplement standard travel diary efforts are doing so in an effort to identify survey under-reporting and incorporate corrections [54]. The objective of this chapter is to develop a new methodology to automate the identification of the activity type for passively collected GPS data, which would increase usefulness of GPS-based travel data for use in model building.

Previous research efforts to identify the trip purpose based on the trip ends, required an underlying layer of the land use type [16, 26, 27, 55, 56]. However, it is difficult to find high quality geographically referenced land use data, as mentioned by both Wolf and Schönfelder in their studies [16, 26]. The availability of accurate land-use parcel data also limits the boundary of space within which the trip purpose can be identified.

Commercial mapping software, such as Microsoft's MapPoint, include geo-coded business locations and points of interest. Because these commercial software applications cover the entire United States, there is no difference in the land use format from once city or region to another. Hence, a single format of data for the entire US from the commercial software helps in the automation of land use search procedures, irrespective of the city [12]. This chapter explores the use of a standalone version of the MapPoint software, coupled with Perl scripts, to identify potential trip purpose and activity based upon proximal land use characteristics at the trip end.

Activity Types

Traditional travel diaries are used to obtain detailed information on the type of activity undertaken for each trip made, such as visiting a grocery store or a service. However, it is not possible to impute the activity type to a very fine level with automated methods because different types of businesses, such as services (banks, gas stations etc.) are likely to be very close or sometimes even within other businesses. For example, bank teller windows may be inside grocery stores, and medical clinics may be located within pharmacies. Therefore, it is necessary to aggregate activities that have similar impact on travel decisions into a single type so that the activity types can be translated meaningfully into travel demand modeling.

Activities based on home, work and school purposes are unique and can be easily identified within a GPS data stream [12]. Travel related to picking up or dropping off someone are unique and have different characteristics. Activities at grocery stores, department stores, and services are necessary activities for the maintenance or sustenance of the households. Activities at theaters, recreational parks, visiting friends and family, stadium etc. can be flexible and be undertaken at the discretion of the households. It is also possible that people travel to mixed-use developments for both sustenance activities and discretionary activities; hence, multipurpose activity type should be identified. Restaurants can be visited for sustenance and/or for socializing with friends and family. The dissertation will assume that most visits to restaurants are for sustenance such as lunch or breakfast rather than for socializing, and are therefore grouped with Maintenance activities. Table 1 shows all the type of places that are available in MapPoint and their proposed activity type classification.

Table 4.1 Cross Table between MapPoint Place Type and Activity Type

MapPoint Place Type	Activity Type	MapPoint Place Type	Activity Type
Airports – Major	Pickup-Drop Off	Restaurants – Greek	Maintenance
Airports – Minor	Pickup-Drop Off	Restaurants - Indian	Maintenance
ATMs	Maintenance	Restaurants - Italian	Maintenance
Auto Services	Maintenance	Restaurants - Japanese	Maintenance
Bus Stations	Pickup-Drop Off	Restaurants - Mexican	Maintenance
Campgrounds	Discretionary	Restaurants - Other	Maintenance
Cinemas	Discretionary	Restaurants - Pizza	Maintenance
Convention Centers	Discretionary	Restaurants - Seafood	Maintenance
Galleries	Discretionary	Restaurants - Steak	Maintenance
Gas Stations	Maintenance	Restaurants - Thai	Maintenance
Hospitals	Maintenance	Schools	School – Daycare
Hotels and Motels	Discretionary	Shopping	Maintenance
Landmarks	Discretionary	Casinos	Discretionary
Libraries	Maintenance	Stadiums and Arenas	Discretionary
Marinas	Discretionary	Subway Stations	Pickup-Drop Off
Museums	Discretionary	Theaters	Discretionary
Nightclubs and Taverns	Discretionary	Train Stations	Pickup-Drop Off
Park and Rides	Discretionary	Banks	Maintenance
Police Stations	Maintenance	Grocery Stores	Maintenance
Rental Car Agencies	Maintenance	Ski Resorts	Discretionary

Table 4.1 continued

MapPoint Place Type	Activity Type	MapPoint Place Type	Activity Type
Rest Areas	Discretionary	Golf Courses	Discretionary
Restaurants - Asian	Maintenance	Wineries	Maintenance
Restaurants - BBQ	Maintenance	Amusement Parks	Discretionary
Restaurants – Chinese	Maintenance	Parking	Maintenance
Restaurants - Delis	Maintenance	City/Town Halls	Maintenance
Restaurants – French	Maintenance	Pharmacies	Maintenance
Post Office	Maintenance	Community Centers	Discretionary

Exploration of Variables that Influence Activity in the Minnesota Data

The activity that one wants to undertake dictates potential activity locations. Hence, knowledge of the activity location will definitely help in the activity type imputation process. Intuitively, one also expects the travel distance, duration of activity, day of week and time of day will be correlated with the activity type at the end of a trip. However, the correlation between these factors and the activity type may not be strong enough to effectively use them in the imputation of the activity type. In this section, the travel diary data collected from University of Minnesota I-35 Bridge Study will be explored to better understand the relationship between the activity type and these factors. A classification tree analysis will be employed to explore the relationship between activity type and travel distance, time of day, day of week, and activity duration.

Standard predictive models, such as linear regression models, are global models which have a single predictive formula designed to represent the entire data space [57]. However, when data have numerous features that interact in complicated non-linear ways, assembling a single global model may not effectively represent the data space. In such situations, partitioning or sub-dividing the data space into smaller regions where the interactions are manageable can be an effective solution. Classification tree or regression tree recursively partitions data into smaller chunks of data space that can be fitted with simple models. The global model has thus two parts, the recursive portioning into cells and simple fit for the data in the cells.

Classification trees help to quickly assess data since the tree method highlights the important variables that affect variability in the data. Classification trees use the mode of the dependent variable if the dependent variable is a categorical variable and the mean if it is a continuous variable.

Figure 4.1 shows the classification tree that includes all the revealed activities by the participants of the University of Minnesota study. In the classification tree at each node, the predominant activity type is presented. When a condition at the node is satisfied, the branch goes to the left and when the condition is not satisfied, the branch goes to the right. The first significant variable to affect activity is time of day. Trips starting before 10AM are more likely to be work trips and Home trips dominate the rest of the day. Maintenance trips are more likely in the afternoon after 2PM. Due to the large number of Home, Work and Maintenance trips, the classification tree model is biased by the frequencies and the activities with smaller frequencies do not show on the classification tree.

To explore other activity types that do not have significant frequencies, a second classification tree was created using weights that were inverse to the frequency of each activity type. The results of the classification tree are presented in Figure 4.2 and the counts now represent the fraction of trips with those activity types. The first split is still influenced by time of day with AM trips pre-dominantly being work trips and PM trips being Home trips as expected. Along the work trips branch the next split is by day of week with the weekends being different from weekdays as expected.

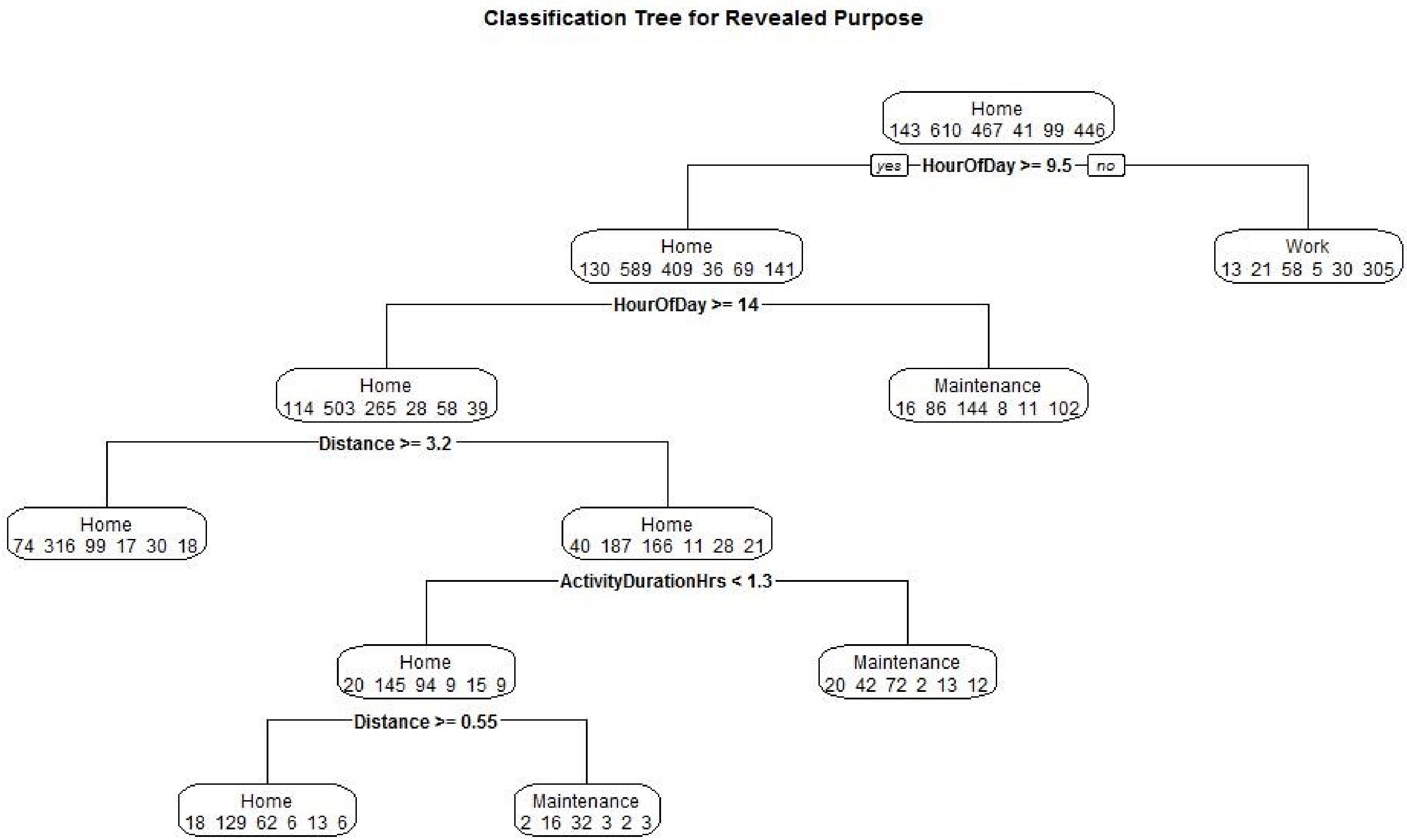


Figure 4.1 Classification Tree of Revealed Activity including all Activities

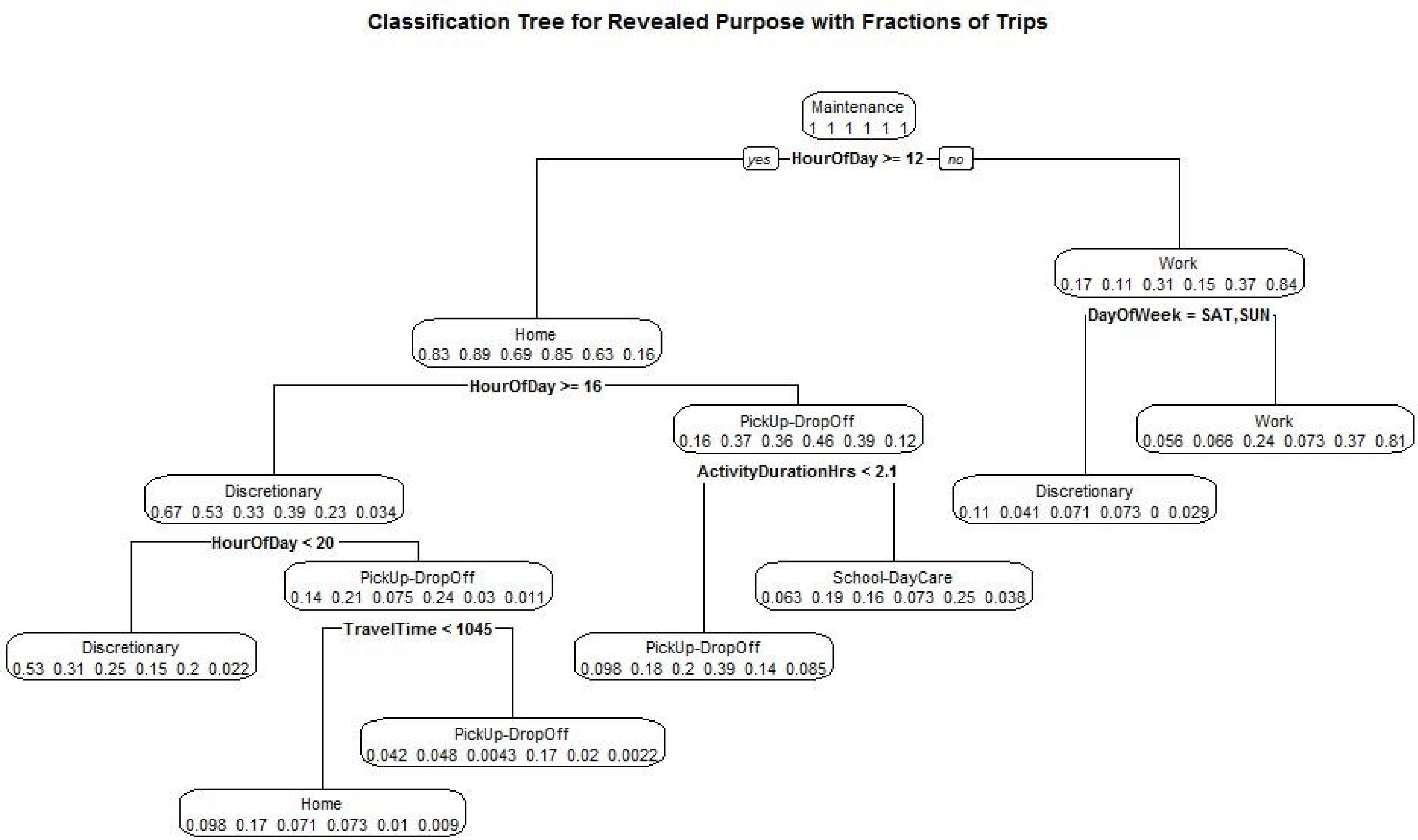


Figure 4.2 Classification Tree of Revealed Activity with Inverse Frequency Weighting

On the PM side of the branch, the next split is by time of day with a significant number of discretionary trips happening after 4PM. Pick-up and DropOffs appear to happen more frequently between 12PM and 4PM. The levels below that show further splits, but the fraction of trips that the leaves represent are very small and the inferences from those leaves are unlikely to be meaningful.

From the classification tree analysis, time of day and day of week have good correlation with the revealed trip purpose and the other variables may not significantly impact the revealed trip purpose. The time of day variable needs to be further explored. Figure 4.3 shows a strip chart of the revealed activity type by the hour of day. The activity types are along the horizontal axis and the hour of day is along the vertical axis. Each ‘o’ symbol represents a trip and the symbols within the same category overlap when many trips are conducted at that time. Therefore, every hour that has significant activity types has darker shades with the overlapping ‘o’s. The discretionary trips occur predominantly in the evenings, and the trips to home are spread throughout the day with more trips in the afternoon. The maintenance trips are spread throughout the day. The pick-up and drop-off trips are very few and they are spread out throughout the day. The school-daycare trips are clustered in the morning and afternoons. The work trips are clustered in the morning.

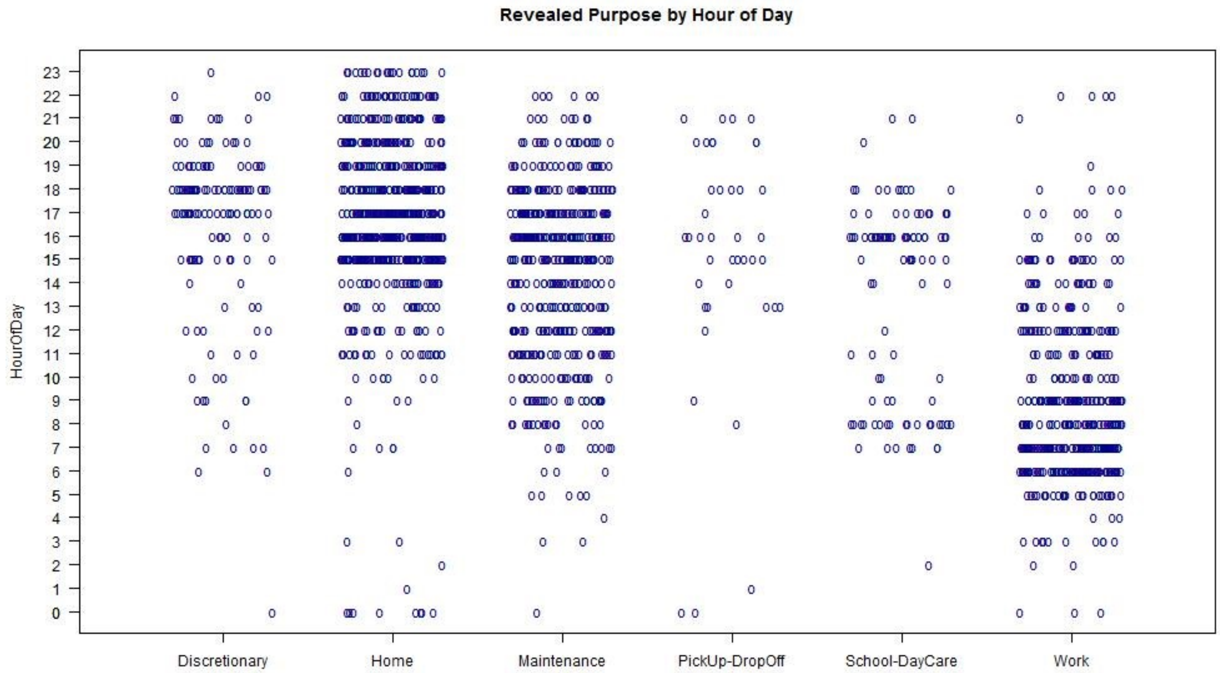


Figure 4.3 Revealed Purpose by Hour of Day

While Figure 4.3 reveals the characteristics of activity types by time of day, the activity types are spread out throughout the day even if it is a smaller percentage. Figure 4.4 further explores the school activity by hour and it clearly shows two clusters, with very few trips happening after 7PM. Based on this analysis, the time of day should be considered especially for trips that occur at school locations. The discretionary trips are clearly clustered towards the evening. However, 17% of discretionary trips happening before noon and therefore the time of day cannot be the only variable used in identifying the discretionary activities.

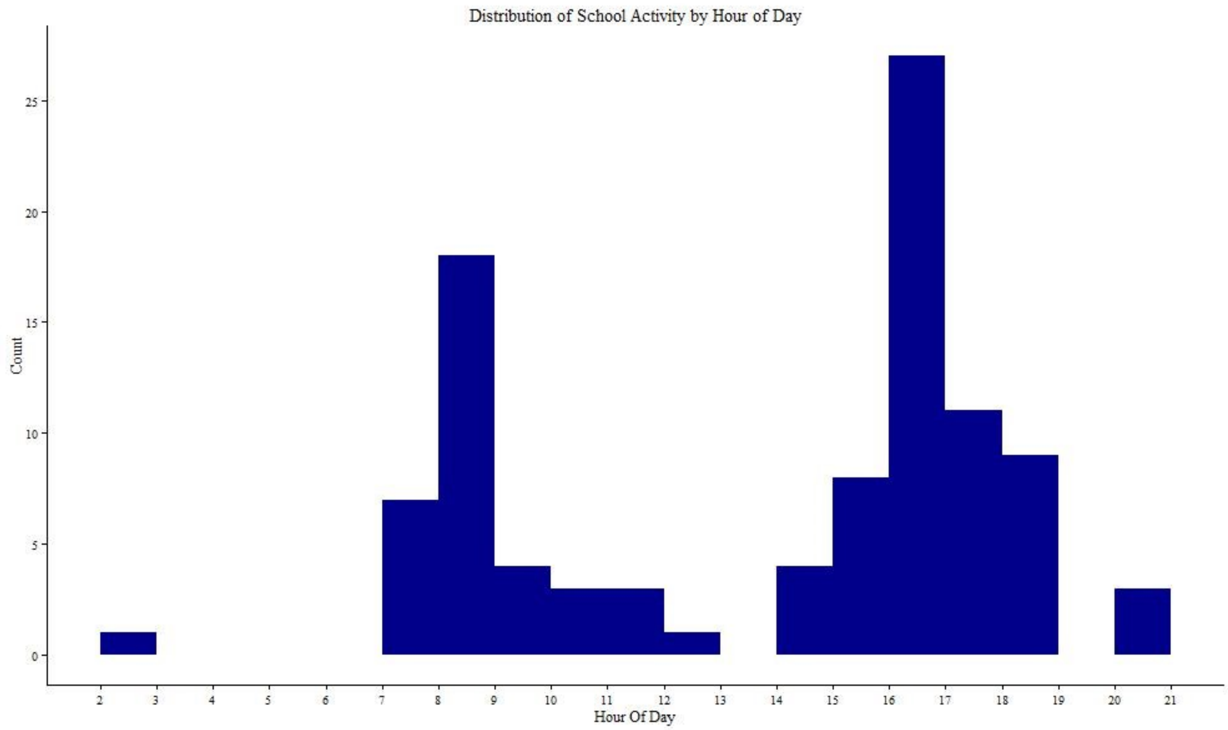


Figure 4.4 Distribution of School Activity by Hour of Day

Assumptions

To develop a methodology to identify the activity at the end of a trip, the following assumptions are proposed based upon the exploration of the University of Minnesota I35 Bridge study data:

- The radius within which people tend to park their vehicles and walk to a destination is 0.3 miles. Most locations in the US are accessible by vehicles and people tend to park as close to their destination as possible. The radius of search is assumed to be 0.3 miles to accommodate locations with large parking lots.
- The activity locations that are closest to the trip end are the most likely locations visited by the individual. It is likely that a series of exceptions should apply to this assumption (depending upon land use configuration, mixed use composition, etc.). But for the purposes

of this analysis, the closest location was examined first. For example, three locations are returned by the search that are within 0.3 miles of the trip end and two of them are located 0.05 miles and the third is located at 0.07 miles . The methodology will only consider the two locations that are 0.05 miles away and not the third.

- The search radius for home location is 500 feet from the trip end. Vehicles are parked as close to the home location as possible, hence the radius of search is tighter. Preliminary analysis indicates that larger radii may be required for apartment dwellings.
- The search radius for work and school locations are 1,000 feet from the trip end. Parking lots at work and school are frequently much further from the office or school location.
- If no businesses or points of interest within 0.3 mile of the trip end, it is possible that the individual stopped at an unlisted business or at a residential neighborhood. Since there is no way of finding the purpose, these locations will be classified as 'Unknown' purpose.
- 'Potential MultiPurpose' activity type implies that there are more than one of the activity types available at the trip-end. When there are multiple activity types close to the trip end (e.g. a trip to a regional or strip shopping mall), it is not possible to conclude whether the trip was made for a single purpose or multiple purposes. For this research effort, such trips are coded as MultiPurpose, even though the trip may actually be for a single purpose. However, it was observed that surrounding the school locations other locations such as restaurants, points of interests and stadiums occur that are part of the school. Therefore, if a school location is one of the location types, the location will be assigned as school.
- Restaurant visits can be sustenance such as lunch or breakfast or discretionary when it has a social element. Based on revealed purpose data of the University of Minnesota study few of the restaurant visits are for social purposes and mostly it is for sustenance purpose. The visit

to a restaurant will happen even if there was no social company since the individual needs to eat. Therefore the research will assume that all visits to restaurants as maintenance activity.

- Activities at school also include discretionary activities such as going to a stadium or attending a cultural event. The activities at school locations after 7 pm and during weekends will be assumed as discretionary activities since most of these activities occur during non-school hours.

Methodology

The first step is to process the raw GPS points to identify trips, trip ends, trip duration, trip distance, start timestamp, end timestamp, and eliminate bad GPS points as detailed in Chapter 3 of this dissertation. The next step is to geo-code the home, work and school locations. Standard household demographic data and address information for home, school, and work locations are usually collected during participant recruitment. It has been observed that during longitudinal surveys, participants change household and work locations, meaning that follow-up surveys in longitudinal efforts are required [48]. The geo-coded work location may not be where the participant is parking their vehicles. To ensure spatial accuracy, the work locations and the home locations need to be verified using all of the longitudinal travel data that are collected. The home location can typically be identified as the most frequent trip end of all trips that occur between 6:00 PM and 6:00 AM. The work location(s) can typically be identified as the most frequent trip end of all trips that occur between 6:00AM and 10:00AM. Frequently households have multiple work locations and the vehicle can travel to either location. For the University of Minnesota study only one participant per household who commutes across the I-35W bridge were monitored. Based on heuristic analysis of the trip data a second location which has characteristics of a work location such as trips in the morning peak and consistent week day repetitions were identified. These location can possibly be the work location for a second household member. The second location is also assumed as a work location if the frequency of that location is at-least 10 over a four month period.

The first step in activity identification for a trip end is to find its distance from the home, and work locations for that household. If the distance falls within the search radius, the trip purpose is assigned either to Home or Work. If the trip end is not Home, or Work, then all businesses within 0.3 miles of the trip end are identified. The algorithms consider only the businesses/places of interest from the search results that are closest to the trip end and find the place type classification of MapPoint for these locations. Using the cross classification table shown previously in Table 4.1, an activity type is then assigned to the location. If all of the places under consideration are of the same activity type, then that activity type is assigned to the trip end. If a school location is found, then the time of activity is considered. School activities starting after 6 PM are classified as discretionary as explained in the assumptions. If there is more than one activity type, the 'Potential MultiPurpose' activity type is assigned to the trip end. However, if there is a school location among the location types, and the hour of day is earlier than 6 PM, then the school activity is attributed to that location. If there are no businesses/places of interest within 0.3 miles of the trip end, assign 'Unknown' activity type to the trip end. A flow chart illustrating the script logic for the algorithms described above is provided in Figure 4.5.

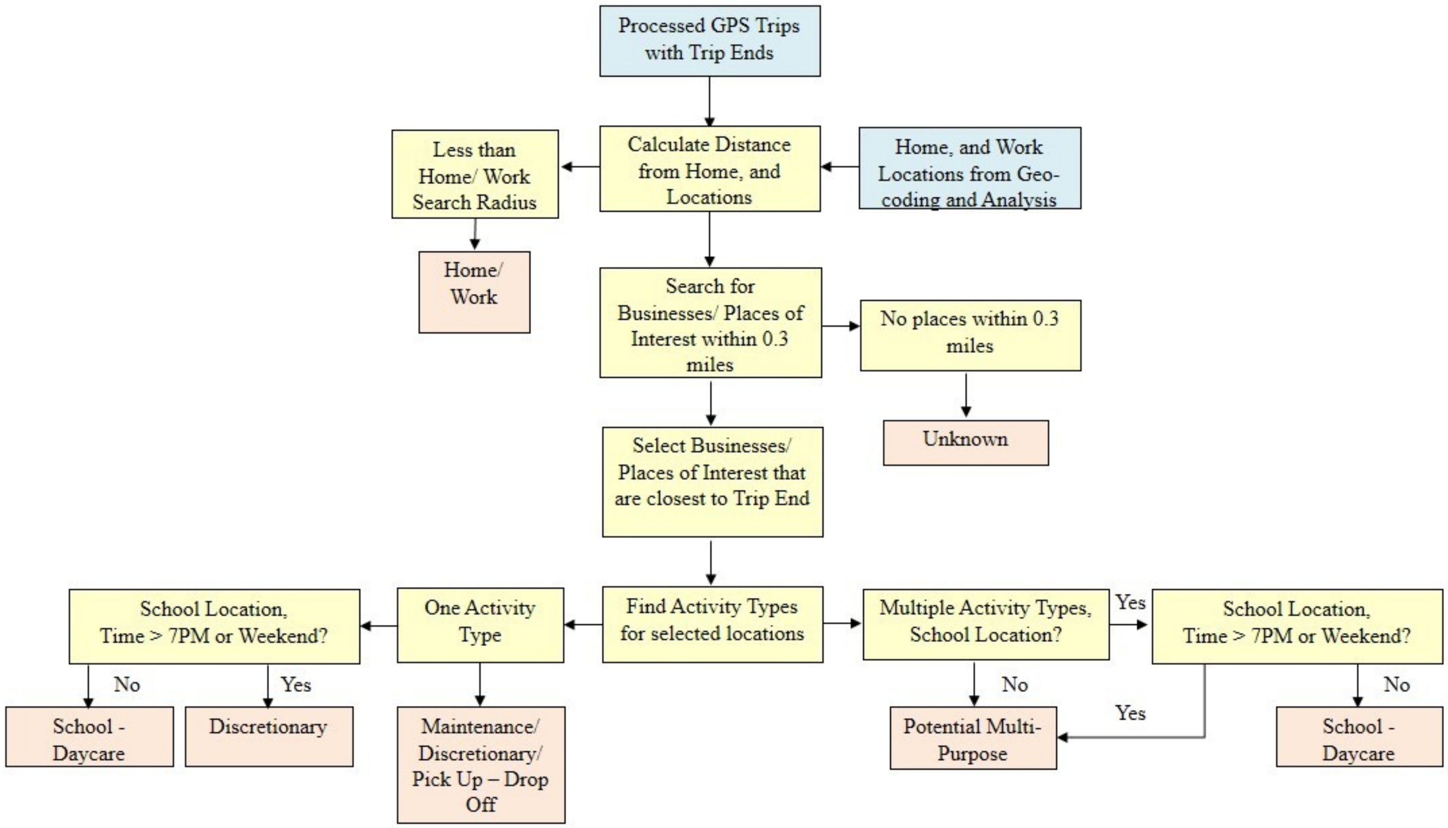


Figure 4.5 Flow Chart of Activity Identification.

A series of Perl scripts were developed to process trip end coordinate data and to identify the revealed trip purposes from the raw data. The Perl scripts also automated the methodology so that it can be readily applied to large data sets.

Summary

A new methodology that automatically identifies the proposed travel activities for passively collected GPS data has been proposed. The proposed methodology does not require human investigation of the GPS data to identify the activity type. This methodology uses commercial mapping software, such as MapPoint, in the place of geographically referenced land use data. This helps make the methodology applicable anywhere in the United Sates and eliminates the variability in the data formats of the land use data by different organizations. The various assumptions that go into the methodology are based on passively collected data from instrumented vehicles (and the high-levels of contiguous data and spatial accuracy associated with vehicle-based data stream). Hence, these assumptions should be re-evaluated if data are collected from hand-held GPS loggers (which are generally not as accurate as in-vehicle logger with a fixed antenna) or by other means. The next chapter will evaluate the methodology and apply it to the Commute Atlanta data set.

CHAPTER 5

ACTIVITY TYPE IDENTIFICATION METHODOLOGY – APPLICATION AND EVALUATION

The activity type identification methodology developed in Chapter 4 are applied and evaluated in this chapter. The first section describes the application of the activity type identification methodology to the University of Minnesota I-35 Bridge study data and the evaluation of the methodology. The next section discusses the limitations of the methodology. The methodology needed to be modified to conform with different assumptions applicable to the Commute Atlanta dataset (described in the third section). The results of the activity type identification for the Commute Atlanta data set are then be presented in the final section of this chapter.

Case Study

The case study using the data from the University of Minnesota Travel Survey was undertaken to evaluate the methodology developed in Chapter 4. As part of this analysis, scripts were developed in Perl that would implement the activity identification methodology on the GPS data. This section compares the trip purpose results calculated using this methodology with the trip purpose revealed from the travel diaries.

The trip purpose data provided by the participants was at a disaggregate level. For example, participants reported fast-food dining as an individual category. For the purposes of the automated trip purpose comparisons, these data were first aggregated into the general trip purpose categories of Home, Work, Maintenance, and Discretionary activity types as described in Chapter 4. If the participant reported multiple activity types, then the activity is assigned 'Multipurpose' as against "Potential Multipurpose" in the calculated activity.

The research team also found that about nine percent of the trips had purpose coded as "Other". After examining a random subset of these trips, the research team believes that when a participant could not recall their trip purpose they coded it as "Other". For the purposes of this comparative analysis, the approximately 227 trips (9 %) recorded by participants as 'Other' were eliminated from the analysis. One household was also using their vehicle for commercial purpose and was excluded from the study. Upon detailed analysis of the data stream, it also appears that one household may not have taken trip purpose reporting seriously, as evidenced by random assignment of trip purposes to known home and work locations. This household had completed the travel diary for almost every day the vehicle was instrumented and the recorded purposes were random. Hence that household was eliminated from the analysis. The final dataset has 1806 trips (82.7% of original data).

Figure 5.1 illustrates the distribution of the travel diary activities by the participants and Figure 5.2 provides the distribution of the calculated activity by the automated MapPoint methodology. From Figure 5.1 and 5.2 we can see that there are more maintenance activities in the calculated activity distribution compared to the travel diary activity distribution. About 29 trips (1.6%) fall under the 'Unknown' category. Those trips may have ended in residential neighborhoods for social visits or are near businesses or other locations not listed in the MapPoint database.

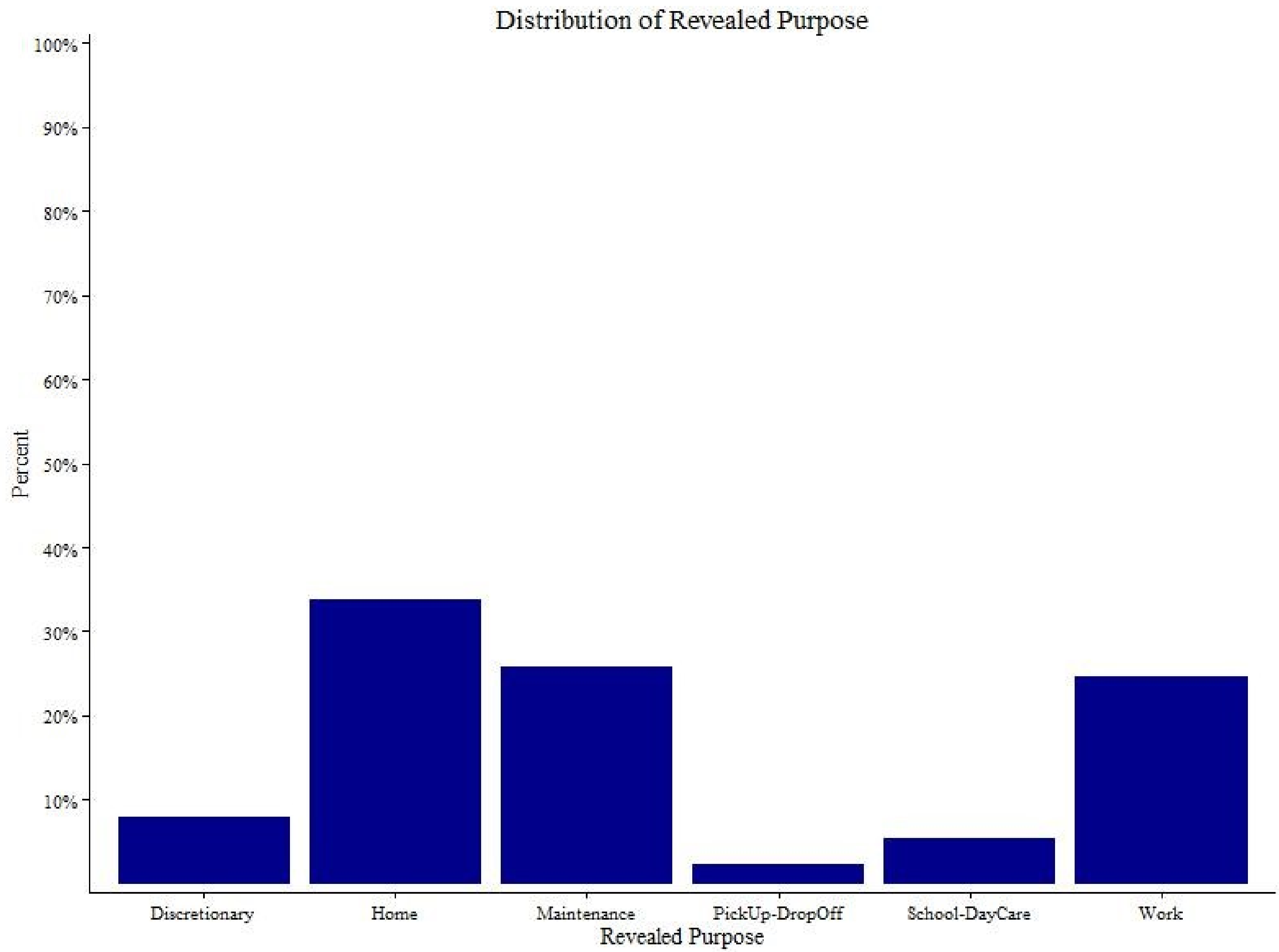


Figure 5.1 Distribution of Travel Diary Activity (n=1806)

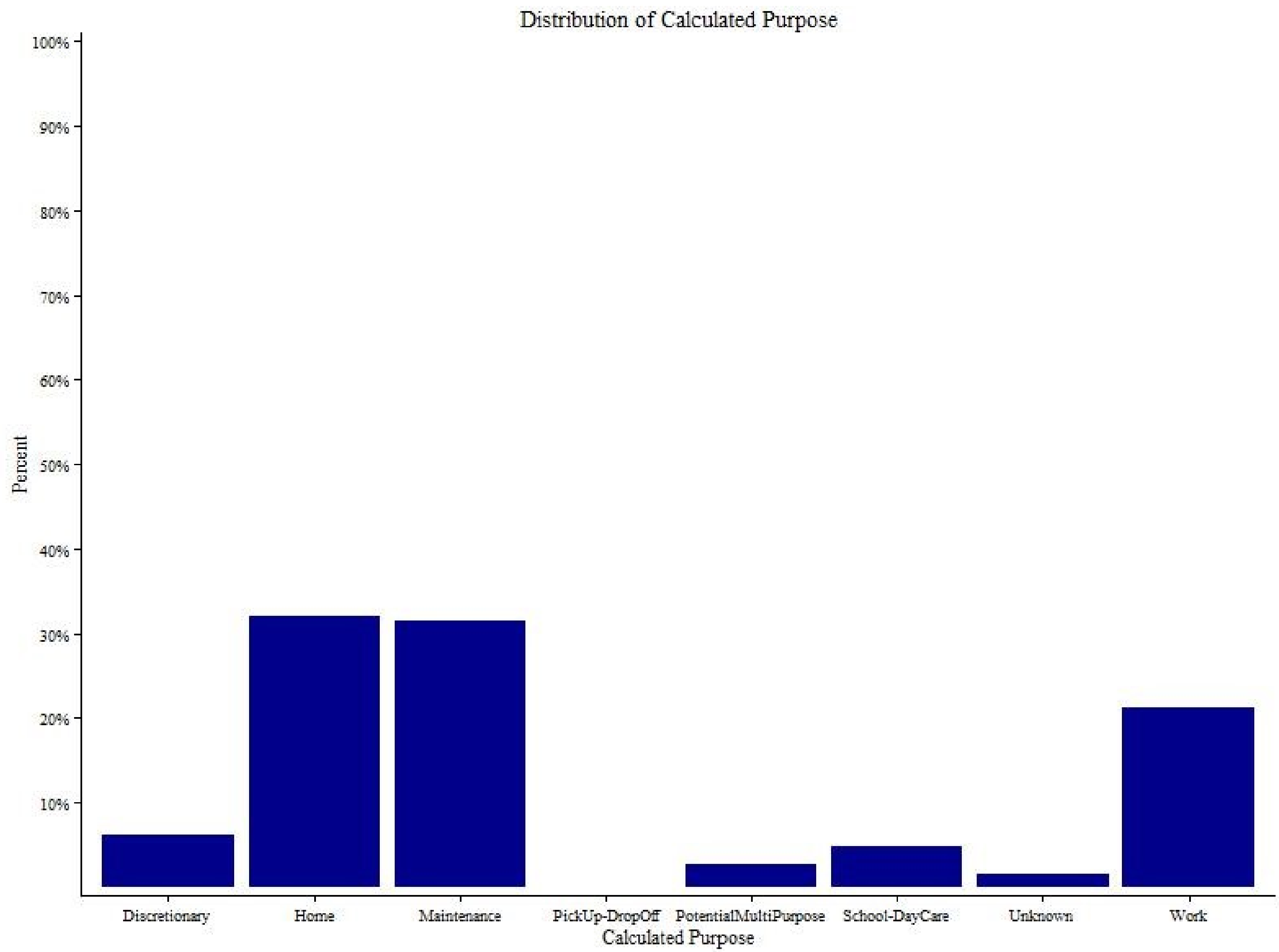


Figure 5.2 Distribution of Calculated Activity (n=1806)

Figure 5.3 shows the bar chart of the distribution of the calculated activity by each travel diary activity. Figure 5.4 shows the bar chart of the distribution of the travel diary activity by each calculated activity. Table 5.1 shows the numerical counts of the Cross-tabulation between reported activity and calculated activity. Travel Diary Home activities are accurately calculated in 82.1% of the cases, travel diary Maintenance activities are identified with 69.4% accuracy, and travel diary Work activities are identified with 71.7% accuracy. However, calculated discretionary activities match with travel diary discretionary activities 27.3% of the time, school and daycare activities match 22.2% and none of the pickup and drop-off activities were matched. Overall, 66.8% of the calculated trip activities match with the travel diary activities.

Table 5.1 Cross-tabulation of Travel Diary Activity vs Calculated Activity

		Calculated Activity								
		Discretionary	**Home**	**Maintenance**	**Pick-Up Drop-Off**	**Potential Multi-Purpose**	**School - Day Care**	**Unknown**	**Work**	**Total**
Diary Activity	**Discretionary**	39	7	67	0	8	10	8	4	143
	Home	10	501	75	1	4	7	3	9	610
	Maintenance	33	36	324	0	25	15	9	25	467
	Pick-Up Drop-Off	8	12	12	0	1	2	3	3	41
	School - Day Care	10	6	38	0	2	22	0	21	99
	Work	11	17	53	0	9	30	6	320	446
	Total	111	579	569	1	49	86	29	382	1806

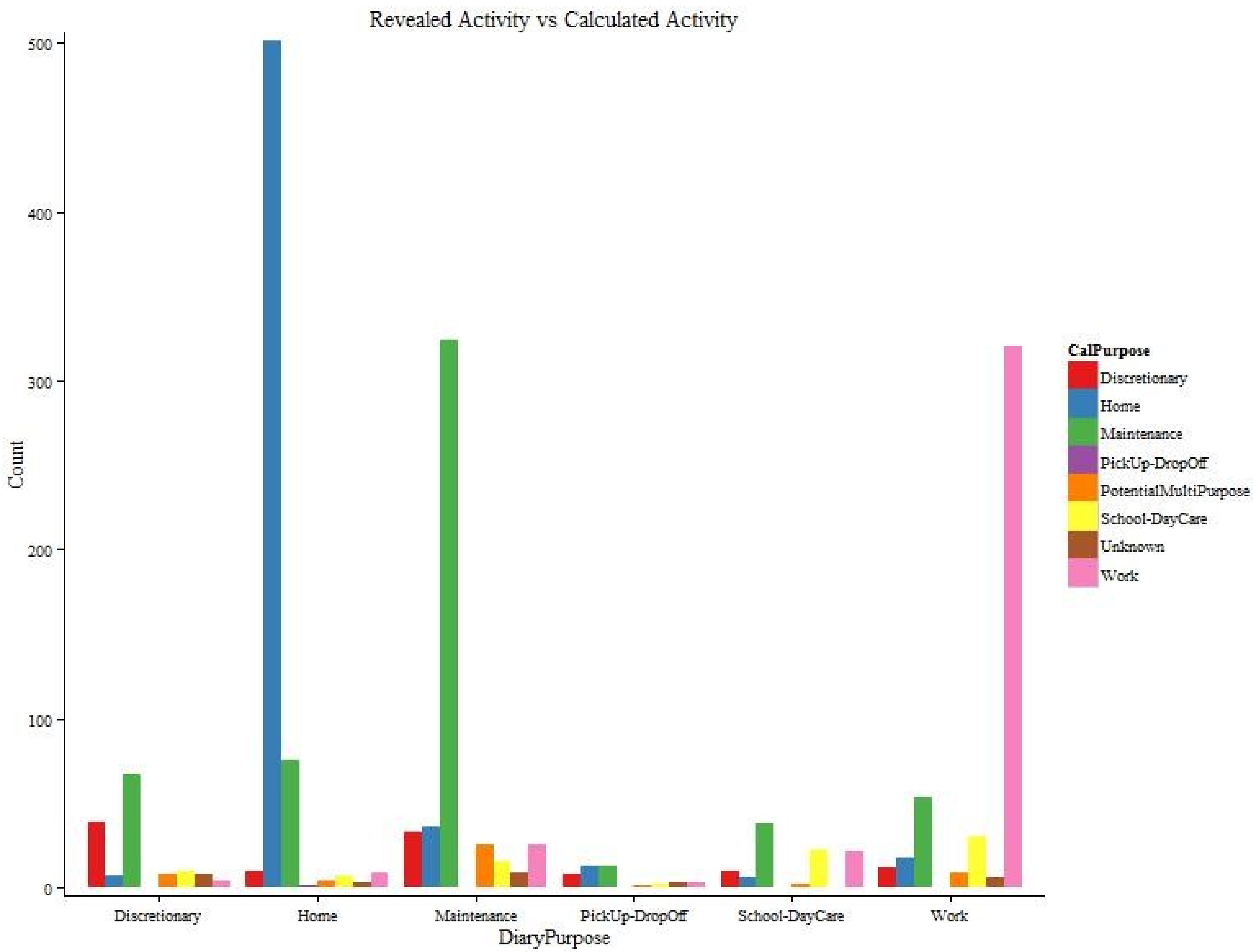


Figure 5.3 Travel Diary Activity vs. Calculated Activity (n=1806)

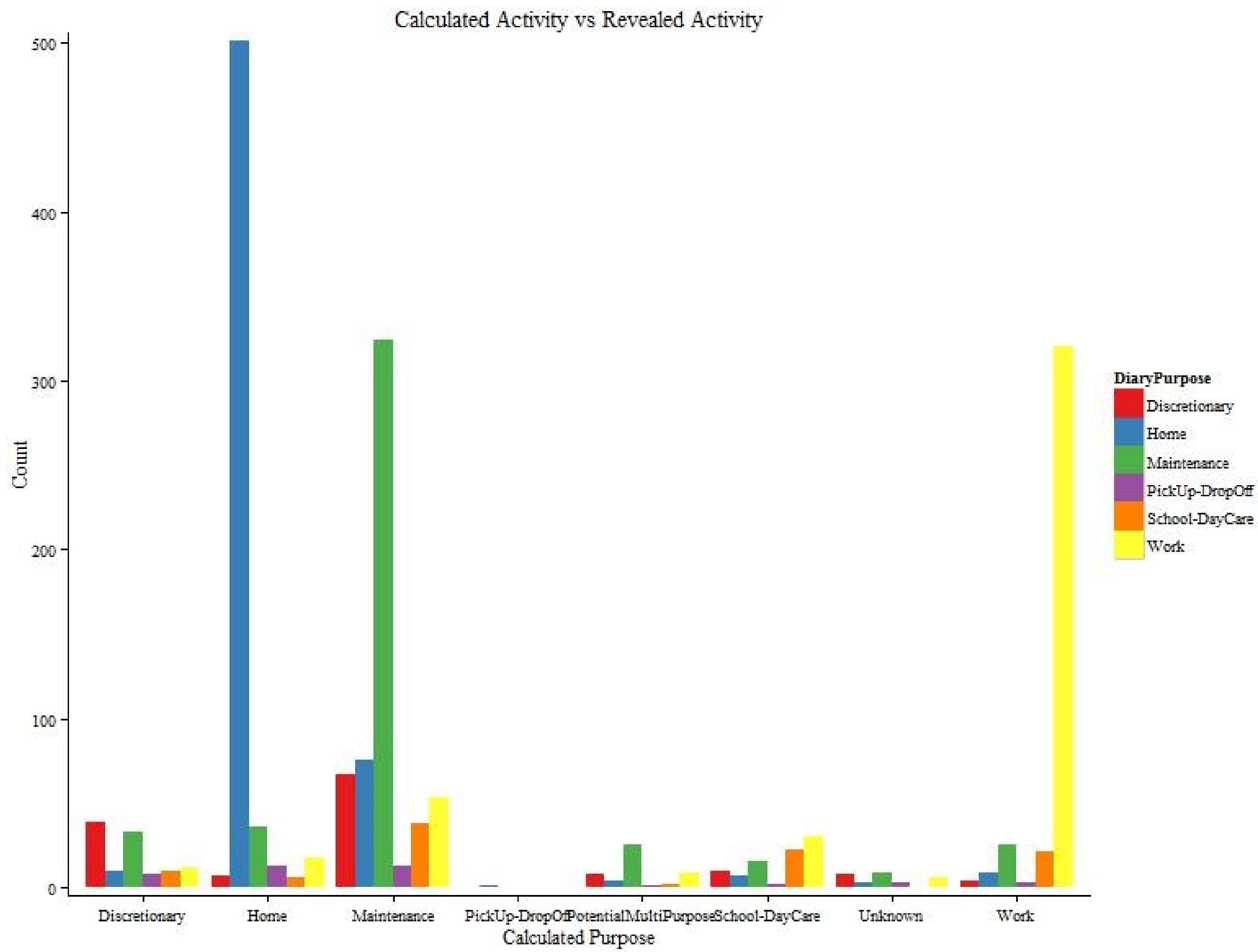


Figure 5.4 Calculated Activity vs. Travel Diary Activity (n=1806)

Discussion

The assumption in this case study is that the revealed travel diary activity for trip purpose serves as ground truth and that the commercial database contains accurate information on all businesses surrounding a trip end. However, upon close examination neither assumption is always true. Analyzing travel diary trip purpose along with the GPS traces and time of day, not all trip purposes are accurately coded. For example, one participant coded three consecutive trips starting at 16:04, 16:18 and 16:50 as trips to home. The first two trips ended at-least a linear mile away from the home location and the last trip was the one that ended at home. In this case, the participant has obviously coded the first two trips incorrectly. To further understand the reasons for mismatch between travel diary and calculated activity types, the following section explores the mismatched activities.

If there were no businesses close to a trip end location, the activity type is classified as 'Unknown'. On exploring the 29 calculated trip ends that were classified as Unknown activity type, 19 trips were to residential neighborhoods, 6 of them at discretionary activity locations, 3 at maintenance activity locations and one of the trip ends had bad GPS data. Twenty five of the 29 activities (86%) belonged to the discretionary activity type. To use passively collected GPS data in any demand modeling effort, it will be necessary to identify activity types for as many trip ends as possible. Therefore, this dissertation will assume that trip ends that do not have any businesses near them are residential locations and hence will assign discretionary activity type for those trip ends.

Home and Work Activities

Home and Work activities are calculated using the distance from the known Home and Work locations. Therefore to explore the mismatches for these travel diary activity types, distance from the Home and Work locations are explored. Out of the 109 travel diary Home activities that did not match with the calculated activity type, the closest trip end to the home location was more than 900 feet away and the farthest was 247 miles away. Figure 5.5 shows the distribution of the distance between those activity locations and the Home locations. It appears that for those activities, they were either marked as Home activities in error or there is a second home location such as the home of a significant other that is not identified in the data set.

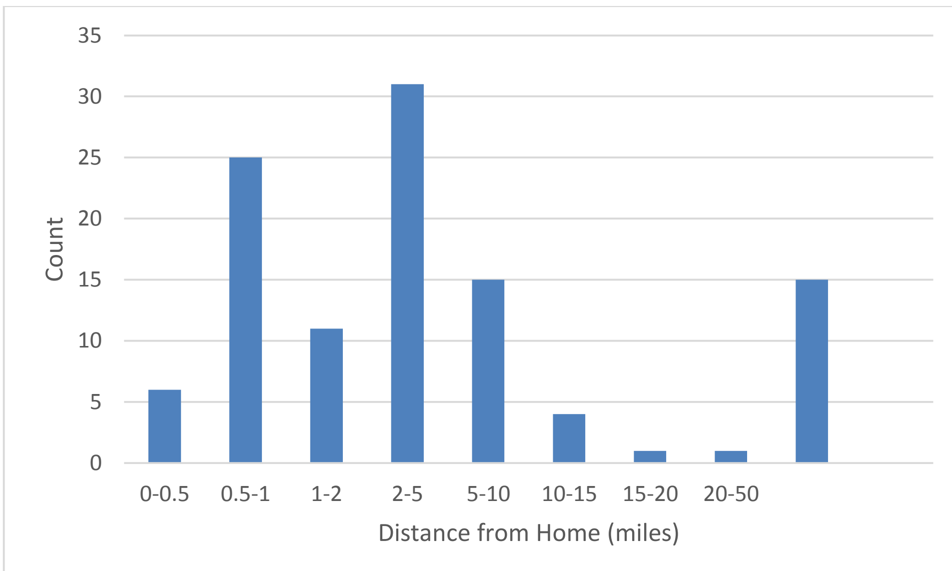


Figure 5.5 Distribution of Distance (miles) from Home Location of non-matching Travel Diary Home Activities

Among the 446 travel diary Work activities, 126 of the activities did not match with the calculated activities. Figure 5.6 shows the distribution of the distance between activity location and work location for the 126 activities. Most of the un-matched travel diary Work activities occurred 2 to 10 miles from the known work location. Further evaluation of the 126 activity locations indicates that 18 of those 126 activity locations were gas stations or auto service centers, identifying a potential for misidentification of work trips. Also, 30 of the 126 mismatched travel diary Work activity locations were near a school, which may either be a school drop-off trip or an unidentified work location. The 30 trips that were identified as work activities in the travel diary data, but were calculated to be school activities, were made by 10 different participants, implying that it is more likely to be an error in reporting the purpose than being an unidentified work location.

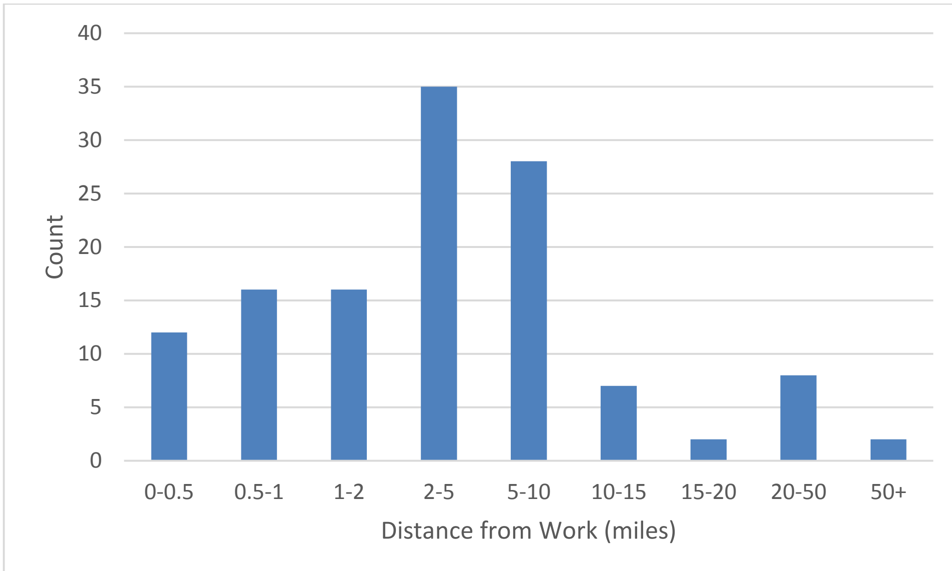


Figure 5.6 Distribution of Distance (miles) from Work Location for non-matching Travel Diary Work Activities

Other Activity Types

The maintenance, discretionary and pickup and drop-off activity types are identified using spatial characteristics while school-daycare has the time of day and day of week taken into account to differentiate between discretionary activities that occur at school. Of the 143 travel diary Maintenance activities that mismatched with the calculated activities, 61 of those trip ends were either near a Home or Work location suggesting that those 61 trips were either misreported in the survey or there were maintenance activity locations within close proximity of the home or work location. Another 25 of those trips were classified as Potential-Multipurpose trips because of the mixed use characteristics of the trip end location. The rest of the mismatched activities were identified as other type of the activities due to the available information in the MapPoint software.

The travel diary discretionary activities that did not match with calculated activities (104 trip ends) were mostly classified as maintenance activities (67 trip ends). Of the 67 trip ends, 30 were at restaurants which implies that it was a social dining activity. The social nature of those activities cannot be discerned from spatial and temporal data. Also, 13 of the travel diary discretionary trip ends were at an ATM or a bank, indicating probable missing spatial information in the commercial mapping database around those trip ends.

The travel diary school-daycare activities matched 22.2% of the calculated school-daycare activities. On exploring the mismatched activity's trip ends (77 trip ends), 38 of them were identified as maintenance trips due to the lack of school or day care information in the MapPoint database. Another 27 of those mismatched trips were calculated as home or work activities, indicating potential errors in reporting the activity type by the participants and/or the presence of day care at work locations. Hence, with proper coding of day care locations, and a clear identification of day care locations in the household recruitment survey, this error can be all but eliminated.

The travel diary pickup-dropoff activities were not identified using the calculation methodology. On evaluating the trip end locations, only one trip was to a subway station and it was misreported as a Home activity by the participant. The revealed pickup and drop-off activities were to all types of locations including Home, Work, and Schools. Figure 5.7 shows the distribution of the activity durations for the pickup and drop-off activities. Most of the pickup and drop-off activities were longer than 15 minutes, indicating that the activities were merely not pickup and drop-off but also involved waiting or other activities. It is not possible to use spatial or temporal characteristics to identify pickup and drop-off activities. Pickup and drop off activities are rarely used in activity participation models even though they play an important role in studying joint activities and interactions between household members.

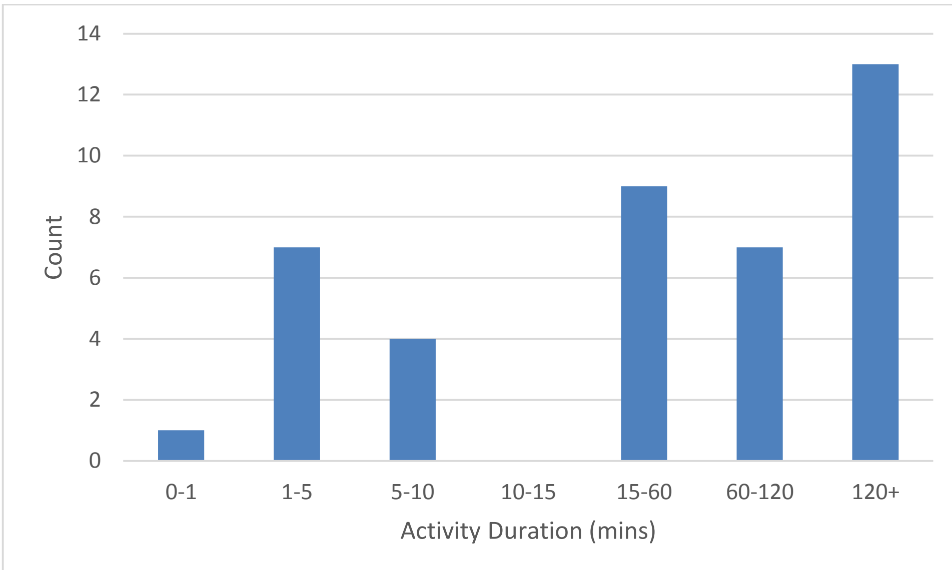


Figure 5.7 Distribution of Activity Duration (minutes) for Travel Diary Pick-up/Drop-Off Activities

Table 5.2 summarizes the unmatched travel trip purpose and the number of trips that were potentially misreported. The potentially misreported home activities were at least more than 0.5 miles away from the actual home location. For the travel diary reported work locations 18 of them were at gas stations and 30 of the locations were at schools as reported in the discussion above. Home or work location that were reported to have maintenance trip activities in the trip diary are potentially misreported. Of the 77 unmatched school locations, 27 of them were at the work location of the participant and are potentially misreported. The misreporting of activities in the travel diary account for at least 41.2% of the trips that could not be matched between the travel diary trip purpose and the calculated trip purpose.

Table 5.2 Summary of Unmatched Travel Trip Purpose with Potential Misreporting

Type	Total Unmatched	Potential Misreported trips	Comments
Home	109	103	Locations more than 0.5 miles away from Home
Work	146	48	30 school and 18 gas stations
Maintenance	143	61	home or work locations
Discretionary	104	0	
School	77	27	work locations
Total	579	239	41.2% of unmatched trips probably misreported

Limitations of the Methodology

It appears that almost 41% of the mismatches in the case study were due to error in reporting the activity by the participant. However, there are some limitations in the developed methodology that limits the accuracy of the calculated activity type. The limitations include quality of the underlying data, incomplete socio-economic information in the data set, the inability to discern between social and sustenance aspects of visits to certain places such as restaurants, the inability to identify the actual activity undertook by an individual when the trip end has mixed-use, and the lack of information of co-participants in the activities.

One of the significant limitations of this methodology is its dependence on the quality of the underlying data. In this methodology we used MapPoint 2010, a commercially available software that is available in the same format across the United States. While the MapPoint data contain comprehensive information on businesses and points of interest, a few location types (such as daycare locations) are incomplete. Using other local data sources such as data from Chamber of Commerce etc. will likely improve accuracy. However, the portability of the methodology will be affected due to the differences in the format and the quality of each local data resource.

Incomplete socio-economic information in the dataset will reduce the accuracy of the calculated activity type. Knowledge of other home (parent home or significant other's home where one might stay), work and school locations of all members of the household will help in accurately identifying those activities. Information on the nature of work and whether the monitored vehicle is used for commercial purposes will improve identification of work activities [11].

The methodology is also limited in identifying work locations for households where work is split into multiple shifts or if the same vehicle is used by different individuals for their commute to work at different times of day. Better collection of socio-economic information during recruitment and on an on-going basis for long term passive data collection efforts would help address this limitation.

The methodology cannot identify drop-off and pick-up activities that occur on the road network. The trip-ends are identified based on complete engine-off or when the vehicle stops off of the road-network. When a vehicle stops on the road network the automatic algorithm cannot differentiate between a stop at a red light or congestion and the stop for pick-up/ drop-off activity. Pick-up and drop-off activities are typically very short and are difficult to identify even if it occurred off the road network. Some bias may be present in this methodology with respect to identifying pick-up and drop-off activities.

While most visits to restaurants and services where for maintenance purposes, there were a few reported activities that had a social component to it. Another example, would be taking someone to a theater for a play practice might be calculated as a discretionary trip by the methodology, whereas the true purpose is to pickup and drop-off. The methodology cannot discern the true purpose of the activity type in these situations. While most of the activities are calculated correctly, some portion will never be calculated accurately. Information on co-participants for certain subsets of activities might help in identifying such activities more accurately.

The other limitation in this methodology is when a trip end is at a mixed-use location. The calculation classifies the activity as “Potential Multipurpose”. However, none of those trip ends in the travel diary data were revealed to be for multipurpose activity. There were not enough trip ends that fell on this mixed use locations to identify patterns in time of day, day of week or activity duration. Hence, the methodology cannot accurately predict the type of activity at mixed use locations. In application of this method, it might be helpful to assume some proportionality of trip purposes as a function of land uses and travel demand by land use for mixed use developments.

Activity Type for Commute Atlanta Data

This section describes the minor changes to the methodology for identifying the activity type in the Commute Atlanta data and reports the results of the activity type identification. The minor changes to the methodology are required because of the differences between the Commute Atlanta data set and the University of Minnesota I-35 bridge study data set. The Commute Atlanta data is route-processed and trip chains have been identified, whereas the University of Minnesota data were not route-processed. Due to the longer data collection period in the Commute Atlanta data and the availability of all vehicle trips for the 95 households in the analysis data set, the Home, Work, School, Habitual breakfast locations and habitual transit stop locations are more comprehensive and have been deduced [48]. Therefore, these habitual locations that have been identified can be included along with the Home and Work activities.

The Home and Work activity type identification is not changed with Home location having a 500 feet threshold and the Work activity having 1000 feet threshold for identification. The first change that will be made to the proposed methodology in Chapter 4 is to search for trip ends that fall within 500 feet of known habitual locations, such as recurrent breakfast locations, and assign the appropriate activity type to those trip ends. The second change in the methodology is for two households that consistently use the transit to get to work. For those two households, if the trip end is at a known transit stop, during a weekday, and if the activity duration is greater than one hour, the trip end will be assigned Work activity. If not, the trip end at the transit station will be assigned pickup and drop-off activity. The last change to the methodology is assigning the discretionary activity type to trip ends that do not have any businesses within a quarter mile. This change is assumed based on the discussion of the case study evaluation.

Results

The results of the activity identification in the Commute Atlanta data are presented below. Figure 5.8 and Table 5.3 show the distribution of activity types that were identified in the Commute Atlanta Study.

About 2.1% of trips had bad GPS coordinates in the entire data stream and hence the activity type could not be identified for those trips. Home location was 26.2% of the activities, and work location was 8.82% of all activities. Maintenance activities had the highest proportion at 41.4%. Approximately 13% of the activities were classified as discretionary, 4.8% activities were classified as Potential Multipurpose, 3.5% were classified as school and daycare activities, and 0.1% were pickup and drop off activities.

Table 5.3 Percentage of Calculated Activity Type in the Commute Atlanta Dataset

Activity Type	Percentage of Trip Ends
Bad GPS	2.09%
Discretionary	12.99%
Home	26.20%
Maintenance	41.43%
PickUp-DropOff	0.10%
PotentialMultiPurpose	4.87%
School-DayCare	3.50%
Work	8.82%

The next step is to compare this distribution against the national household travel survey to check for potential biases. The distribution of the trip purposes in the 2009 national household travel survey (NHTS) is shown in Table 5.4[58]. The activities identified in the Commute Atlanta dataset are activities that happen at a trip end. In the NHTS data, the identified the purpose of a trip, which implies a commute trip included both trips to work and back to home. The Work trips in the NHTS dataset are almost twice the work activities found in the Commute Atlanta data as expected based on their definitions. School and Church trips, and social/recreational trips in the NHTS data are approximately twice as found in the Commute Atlanta data. However, the maintenance trips are almost the same as both the data sources. This implies that there might be a bias towards identifying maintenance activities using the proposed algorithm.

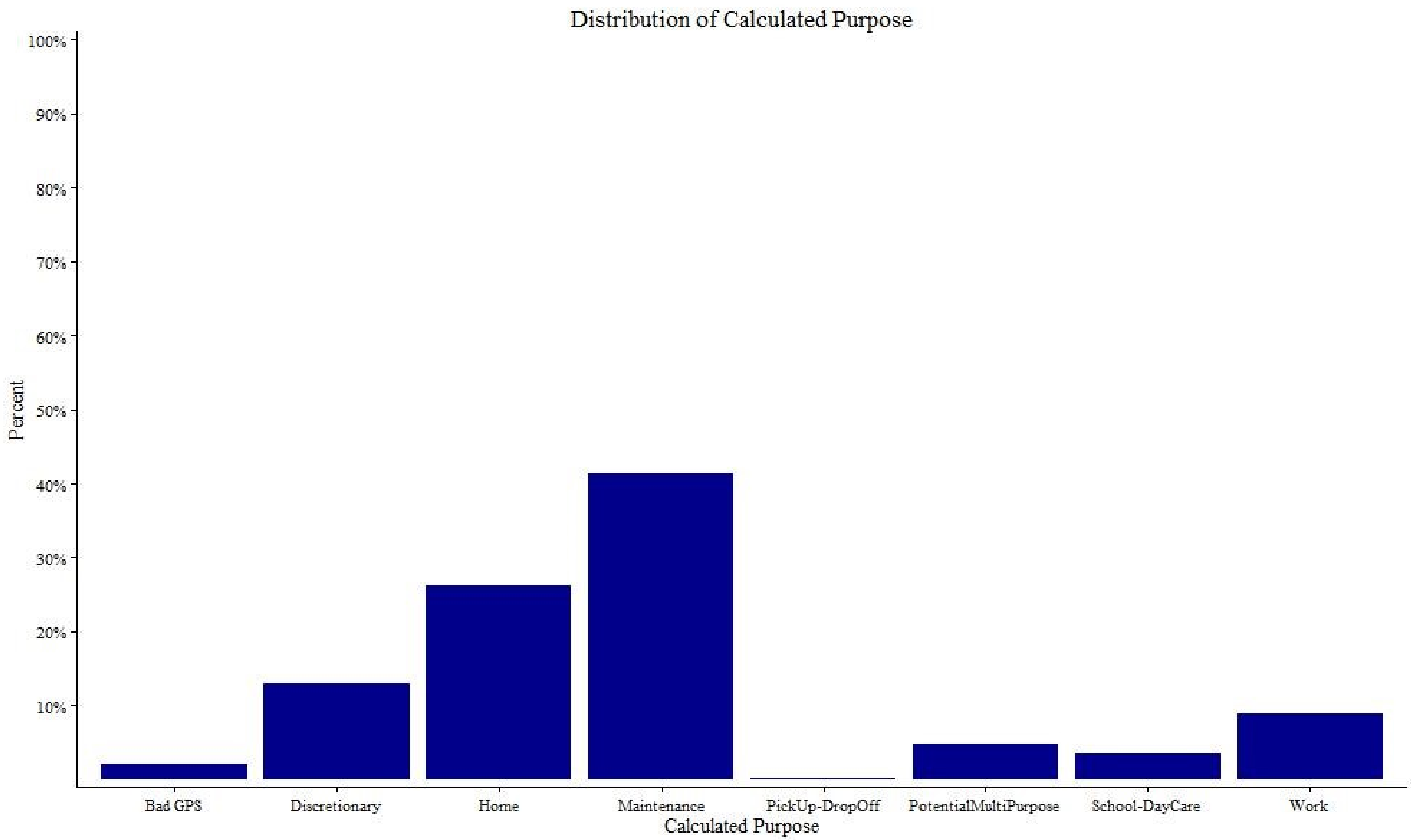


Figure 5.8 Distribution of Calculated Activity Type in the Commute Atlanta Dataset

Table 5.4 Distribution of Trip Purpose from the National Household Travel Survey 2009[58]

Trip Purpose	Women	Men	Average
Work (To and From)	15.7%	21.9%	18.8%
Family and Personal Errands	46.1%	38.7%	42.4%
School/ Church	9.9%	9.4%	9.6%
Social/Recreational	26.8%	28.2%	27.5%
Others	1.6%	1.9%	1.8%

Summary

A case study compared the activity types predicted from this methodology with the travel diary activities. The data collected by the University of Minnesota in 2008 was used for this study. The analysis showed that this methodology can accurately predict Home, Work and Maintenance activities. Using the automated tool to identify discretionary and multi-purpose activities will require significant improvements. Overall the methodology identified 66.7% of the trip purposes conforming to the reveal trip purpose in the travel diary. The methodology may have some bias in identifying certain activities, such as pick-up/ drop-off activity, work activity that occurs during multiple shifts, social visits to shopping or dining etc.

Evaluation of the trips for which the travel diary and calculated activity did not match found that the revealed activity is not always the ground truth. The developed methodology was limited by the quality of the underlying spatial and demographic data. The methodology also was not able to discern when a social component or pickup and drop-off component was part of a trip that ended in other location types. Detailed analysis of the trip purposes that could not match between the travel diary trip purpose and the calculated trip purpose found that at least 41% of the trips were misreported in the travel diary.

Minor changes were made in the methodology for its application to the Commute Atlanta data set to reflect the availability of habitual behavior information in the Commute Atlanta data. The results of the activity type identification in the Commute Atlanta dataset were compared with the distribution of the trip purposes in the national household travel survey data. Work, social/recreational, and school/church trip purposes were twice the activity types identified in the Commute Atlanta dataset as was expected based on how the trip purposes were defined in that survey. The methodology developed may have a bias towards identifying more maintenance activities based on the comparison with the NHTS data. However, due to the lack of ground truth this cannot be explored further.

CHAPTER 6

VARIABILITY IN TRAVEL

As discussed in Chapter 2, travel variability results from the natural daily variation of an individual's transportation needs and desires, but is also affected by feedback from the transportation system itself (previous travel history, previous experience, congestion levels, etc.) and interactions with other parties during travel decision making processes. Traditional travel demand models use cross-sectional data with large sample sizes. Cross-sectional data sets can help explain inter-personal travel variability. However, the travel variability within an individual or household's travel behavior (intra-personal variability) are not sufficiently explained by the traditional models and become part of the error terms in the model [4]. Study of intra-personal variability in travel behavior helps travel demand modelers obtain better analytical results using advanced statistical tools, helps social researchers to better understand travel behavior; and helps policy analysts to obtain better insight into the potential effects of transportation policies over time [2]. While previous research indicates that variability in travel behavior is influenced by socio-economic factors [4, 13, 20, 21, 24, 59], the individual's adventure seeking nature as well as the impact of changes in the environment (such as road closures, or new land use development) also influence the variability [60]. The research reported in this dissertation will use intra-household variability as one of the inputs into the modeling process as a surrogate for the individual's variability-seeking nature on that household's travel behavior.

This chapter presents the different statistical methods to estimate travel behavior variability within a household and describes the methodology that will be used in this dissertation to measure the intra-household travel behavior variability. The first section describes the standard statistical measures of variability. The second section states the assumptions that will be made in this research effort. The third section briefly explores the measures of variability used by other researchers and proposes the methodologies that will be used in the dissertation. The last section summarizes the research efforts presented in this chapter.

Standard Measures of Variability

Standard statistical measures of variability include range, interquartile range, average absolute deviation from the mean, variance and standard deviation, and the coefficient of variation. The potential of each measure for assessing travel behavior variability is discussed in this section.

Range

The range of a sample is an intuitive measure of variability that provides the maximum and minimum values, and the difference between them. The range is very sensitive to outliers and may have different values for two samples with similar dispersion [61]. The range provides valuable information on the extent of the data, but is not the best measure of dispersion of the sample.

Interquartile Range

The interquartile range is another measure that helps avoid the effects of outliers by taking the difference between the first and third quartiles of the sample [61]. The interquartile range is less sensitive to the extreme outliers, but it also ignores half of the data (the highest quarter and the lowest quarter). This leads to loss of valuable variability information in those data and hence may not be the best measure to study travel variability.

Mean Absolute Deviation from the Mean

The mean absolute value of the difference between each data and the sample mean is one measure of variability for day-to-day trip making [59, 61]. The mean absolute value of the difference reflects the magnitudes of the deviations, without the sign of the deviation (to avoid error cancellation). The mean absolute deviation provides a direct comparison between two samples and their dispersions. The mean absolute deviation is also sensitive to outliers [61]. The sample mean absolute deviation is also a biased estimate of the population's mean absolute deviation, i.e. if the sample size is increased, the error between the sample estimate and population value will not necessarily decrease. This variability measure may be used in travel behavior analysis when we have the complete travel data for a household, and will be further explored. The equations below present the mean absolute deviation in mathematical terms.

$$Mean\ Absolute\ Deviation(MAD)_{population=} \frac{\sum_{i=1}^{n}|x_i - \bar{x}|}{n}$$

$$Mean\ Absolute\ Deviation(MAD)_{sample=} \frac{\sum_{i=1}^{n}|x_i - \bar{x}|}{n-1}$$

Where:

x_i i^th data element,

$\bar{x}$ the mean, and

n the number of data points.

Variance and Standard Deviation

An unbiased measure of dispersion is the variance, or the expected value of the squared deviations from the mean. The sign of the deviation is eliminated by squaring the deviations and hence the magnitudes of the deviations define the dispersion measure. The extreme values of the data do affect the variance value, so outliers need to be handled with care [61].

The standard deviation is the square root of the variance and has the same units as the data under consideration and hence easier to interpret. The standard deviation may not be an appropriate measure of dispersion for samples that are skewed [62]. The use of variance and standard deviation in estimating the intra-household variability will be explored in the next section.

The variance and standard deviation are given by the following equations:

$$\sigma^2 = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n}$$

$$s^2 = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n-1}$$

Where:

σ the population standard deviation and σ^2 is the population variance, and

s the sample standard deviation and s^2 is the sample variance.

Coefficient of Variation

The coefficient of variation is a dimensionless measure of the value of the standard deviation, relative to the mean. The coefficient of variation (cv) is the ratio between the standard deviation and the mean of the sample. The coefficient of variation provides valuable information on the dispersion of the sample. If the coefficient of variation is greater than one, the sample is considered to have large dispersion. The magnitude of the coefficient of variation provides the relative variability of the data between two samples irrespective of the means of the samples [61]. Therefore the variability between two samples can be directly compared using their coefficients of variation. The coefficient of variation provides a good measure to study the intra-household travel behavior variability and also compare the measure between households. The coefficient of variation is presented by the following equation:

$$cv = \frac{s_x}{\bar{x}}$$

Where:

cv the coefficient of variation

s_x the standard deviation of the sample

$\bar{x}$ the mean of the sample

Assumptions

The dissertation will make the following assumptions in the analysis of travel behavior variability:

- The trips made by the vehicle reflects the complete travel of the household. While the household will make a small number of walking, biking and transit trips, the commute Atlanta data study did not monitor those trips. In Atlanta 94% of households have at least one vehicle and 87% of all trips made are in vehicles either as a driver or passenger [9, 10]. Walk and School Bus are the other significant modes of transport in Atlanta, accounting for another 10% of the trips [9, 10]. The non-auto trips are mostly undertaken by the households that do not have any vehicles, making the share of auto trips by households with vehicles larger than 87%. The Commute Atlanta study was an instrumented vehicle study and had households that have at least one vehicle. Walk trips and other non-auto trips are useful in planning of sidewalks, bike routes, and transit routes, but may not significantly affect the four step process in identifying new highway and arterial infrastructure needs. While the assumption that vehicle trips constitute the complete travel of the households clearly misses some of the activities undertaken by the households in the Commute Atlanta study, vehicle trips still capture the vast majority of the trips and activities of the households; hence, this assumption will not significantly affect the results from activity participation modeling.

- The number of trips made on each day in a month by a household constitutes the population of that household's travel behavior data set. This assumption extends from the previous assumption that the vehicle trips constitute the full set of trips made by the household. The dissertation work will estimate the intra-household variability in travel behavior for a given month, and the analysis data set includes all the travel data for that month for that household. This assumption leads to the inference that the mean, variance, and other statistical measures calculated for a household month are the population measures and not the sample measures given that all days in the household month are included in calculating the statistical measures.
- The distribution of number of trips, distance traveled and travel time are not normally distributed [11]. Hence, the analytical work will use non-parametric statistical methods for assessing travel behavior variability.

Methodology for Variability in Travel Behavior

Previous research efforts in the literature to assess intra-household and intra-personal travel variability have used variance of number of trips, distance or travel time [4, 13, 20, 24, 63], deviation from the mean [59], variability in route choice [64] and number of new trip ends [3] as measures of variability. Prior research efforts have identified that demographic variables such as income, employment status, household size and the role of the person in the household, also affect intra-personal variability [4, 13, 20].

The dissertation will evaluate the intra-household travel behavior variability in terms of number of trips, daily distance traveled, and daily travel time to identify the variables that effectively reflect the variability-seeking nature of the household in activity participation. The dissertation will apply the mean absolute deviation from the mean, standard deviation, and the coefficient of variation methods and evaluate the most suitable method for determining the intra-household travel variability of each household month. While the intra-household travel behavior variability is expected to reflect the adventure seeking nature of the household and the built environment around the household location, demographic characteristics such as income, vehicle ownership, household size, number of workers, number of students, and number of children are also linked to variability. Therefore, the dissertation will evaluate the demographic variables that have a significant impact on the intra-household travel behavior variability.

Xu found that the number of trips could be represented by a Poisson distribution and daily vehicle miles of travel by a Tweedie distribution [11]. To identify intra-household travel behavior variability, the analyses in this dissertation use the variability in number of trips, daily distance traveled and daily travel time, which are not normally distributed. Therefore, to examine the potential effects of demographic variables and test hypothesis on the intra-household travel behavior variability, the dissertation will use nonparametric statistical methods, such as bootstrap analysis and Mann Whitney U tests. Applying nonparametric methods across the three data sets helps in direct comparison of the results.

Bootstrap

Bootstrap method is a simulation of data that helps with developing statistical inferences such as confidence intervals [65]. The Bootstrap is used to infer the variability of an unknown distribution using large numbers of data sets that are generated by resampling the data with replacement [66]. The bootstrap is based on the premise that in the absence of any information on the distribution of the data, the observed sample contains all the necessary information about the distribution [67].

Suppose there is a data sample X=[X_1, X_2, X_3, ...,X_n] that are drawn independently from the distribution P, and let s(X) be the sample estimate of θ. For statistical inference on θ, the sampling distribution of s(X) is necessary to assess the accuracy and set confidence intervals for the parameter θ. The bootstrap principle replicates the data generating process by sampling from an estimate $\hat{P}$ of the unknown distribution P [67]. The method assumes that X has complete information on the distribution P and can be considered as the empirical distribution $\hat{P}$. The role of the real quantities are then taken over by the analogous quantities in the bootstrapping process. Let $X^* = [X_1^*, X_2^*, X_3^*, \dots, X_n^*]$ is a bootstrap sample from $\hat{P}$

$\theta^* = s(\hat{\mathrm{P}})$ in the bootstrapping process

$\widehat{\theta^*} = \mathrm{s}(X^*)$ is the bootstrap replication of θ

The sampling distribution of $\hat{\theta}$ may then be estimated by its bootstrap equivalent

$$\hat{P}(\hat{\theta} \in A) = P^*(\hat{\theta}^* \in A) \text{ [67].}$$

The dissertation proposes to use the bootstrap to study the mean and confidence bounds of intra-household variability across the socio-economic variables to better understand the effects of the variables on travel behavior.

Mann Whitney U Test

The Mann Whitney U test, also known as the Wilcoxon Rank Sum test, is the nonparametric equivalent of the parametric t-test [68]. The t-test assumes normal distribution and requires the variables to be measured at the interval or ratio scale. Whereas, the Mann Whitney U test is used to test two groups on a single ordinal variable that does not have a specific distribution [68].

Consider two variables X (sample size n_x) and Y(sample size n_y) with continuous cumulative distributions f and g. The variable X will be stochastically smaller than Y if $f(a) > g(a)$ for every a. The Mann Whitney U test is used to test the null hypothesis $f = g$ against the alterative that X is stochastically smaller than Y [69]. To test the null hypothesis, the two samples X and Y are grouped into a single sample and ordered. The elements of the combined ordered sample are then assigned ranks 1 through N where N is the size of the combined sample ($N = n_x + n_y$). If two elements of the sample are tied, either the mean, minimum, or maximum rank of the tied elements can be assigned to the tied elements. Then the sum scores of the rank T_x and T_y within each of the samples X and Y are computed [64]. The U statistic is then given by [68]:

$$\text{If } n_x > n_y : U = T_x - (n_x(n_x + 1))/2$$

$$\text{If } n_x < n_y : U = T_y - (n_y(n_y + 1))/2$$

The U statistic has a discrete or uniform distribution and hence possible to test the null hypothesis with the help of critical values and their associated probabilities [20]. The dissertation will employ Mann Whitney U test to test hypothesis on the effects of the socio-economic characteristics such as income level, presence of children, etc., on intra-household variability.

Spearman's Coefficient

The Spearman's coefficient is the non-parametric equivalent of the Pierson's correlation coefficient. The Spearman's coefficient is a measure of the monotonic association between two variables[70].

Consider two variables U and V and the associated ranking for each of the records as u_i and v_i. The Spearman's coefficient r_s is the product moment correlation coefficient of u_i and v_i and may be computed from the sum of the squared differences [71].

$$S_s = \sum_{i=1}^{n} (u_i - v_i)^2$$

$$r_s = 1 - \frac{6S_s}{(n^3 - n)} \quad [71]$$

Summary

This chapter presents the various methods to quantify intra-household travel behavior variability, which can be used as a surrogate measure for the individual's variability-seeking nature and travel constraints that may be related to the built environment around the individual household. The chapter explored different methods to determine measures of dispersions or variability in samples and population. The range and inter-quartile range were found to be limited in studying travel behavior variability due to the significant influence of outliers on the range and elimination of half the data in the inter-quartile range. The mean absolute deviation, variance, standard deviation and the coefficient of variation are expected to provide better measures of intra-household travel behavior variability.

The chapter stated the assumptions that will be made in determining the intra-household variability. The first assumption was that the vehicle trips constitute the entire travel for a household. The next assumption was that the daily number of trips for each household month constitute the population of the data since all vehicle trips were monitored. The last assumption was that the travel behavior variables such as number of trips, daily distance traveled and the total travel time do not follow normal distributions and therefore nonparametric statistical methods will be used to evaluate intra-household travel behavior variability.

The chapter then described the methodology to determine the intra-household travel behavior variability. The dissertation will use mean absolute deviation, standard deviation and coefficient of variation to measure the variability in number of trips, daily distance traveled and total travel time. The dissertation then discussed nonparametric methods such as bootstrap, Mann Whitney U test, and Spearman's Coefficient that will be used to assess intra-household variability with respect to socio-economic variables.

CHAPTER 7

VARIABILITY IN TRAVEL BEHAVIOR - RESULTS

This chapter presents the results of the travel behavior variability explored in this dissertation. The first section presents the results of variability in distance, travel time and the number of trips and discusses the benefits and limitations of using each of the attributes as a measure of the adventure seeking nature of an individual household. The next section evaluates the measures of variability discussed in Chapter 6 to identify the measure that is most suitable to apply in travel behavior variability. Once a suitable travel behavior attribute and a measure of variability are identified the next section analyzes the relationship between travel behavior variability and the household socio-economic variables. The last section summarizes the findings in this chapter.

Travel Behavior Attributes

The quantitative travel attributes that summarize the average daily travel behavior of a household are number of trips, vehicle distance traveled and vehicle travel time. This section presents the results of the variability estimated for each of the quantitative travel attributes by the three methods proposed in Chapter 6. There are 282,266 trips in the trip dataset, covering more than 1.9 million vehicle travel miles in more than 79,000 hours of travel. Following the results, this section will evaluate the suitability of each attribute's effectiveness in revealing the variability-seeking nature of the individual household.

Number of Trips

The number of trips undertaken is directly correlated to the number of activities outside the home that an individual household seeks. Figure 7.1 shows the distribution of the mean number of daily trips for each household month. The mean of daily trips across household months is about 7.2 trips per day. The mode of the distribution across household months is approximately 6 trips/day. The number of trips follow a Poisson distribution, as noted by Xu [11].

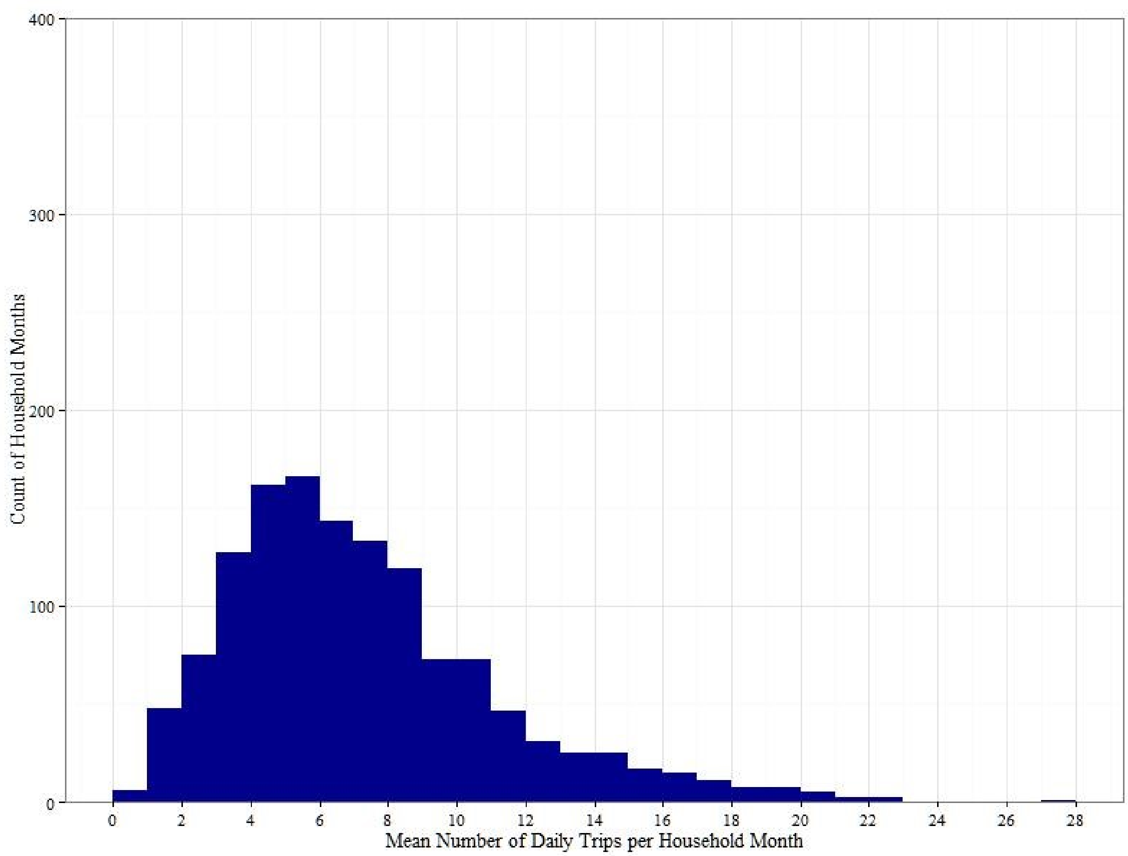


Figure 7.1 Distribution of Mean Number of Daily Trips per Household Month

Figure 7.2 shows the distribution of the mean absolute deviation, standard deviation and coefficient of variation for the number of daily trips per household month. The units of mean absolute deviation and standard deviation are the number of daily trips while the coefficient of variation is dimensionless. The standard deviation is slightly larger than the mean absolute deviation for the number of daily trips since the data points that are farther from the mean have a larger impact on the standard deviation than the mean absolute deviation.

The mean of mean absolute deviation of daily trips is 2.9 daily trips and the mode is 2 daily trips. The mean of standard deviation of number of daily trips is 3.8 daily trips and the mode is approximately 3 daily trips. The mean of the coefficient of variation of the number of daily trips is 0.6 and the mode is also 0.6.

The number of trips is a good measure of the travel behavior of a household because it is directly proportional to the number of activities that an individual undertakes outside the home. In contrast, the duration between trips indicates the activity duration at each location and with increase in number of trips there is less time available to participate at activities. The number of trips variable does not provide information on the spatial extent of a household's activities; hence, that information needs to be determined through spatial analysis.

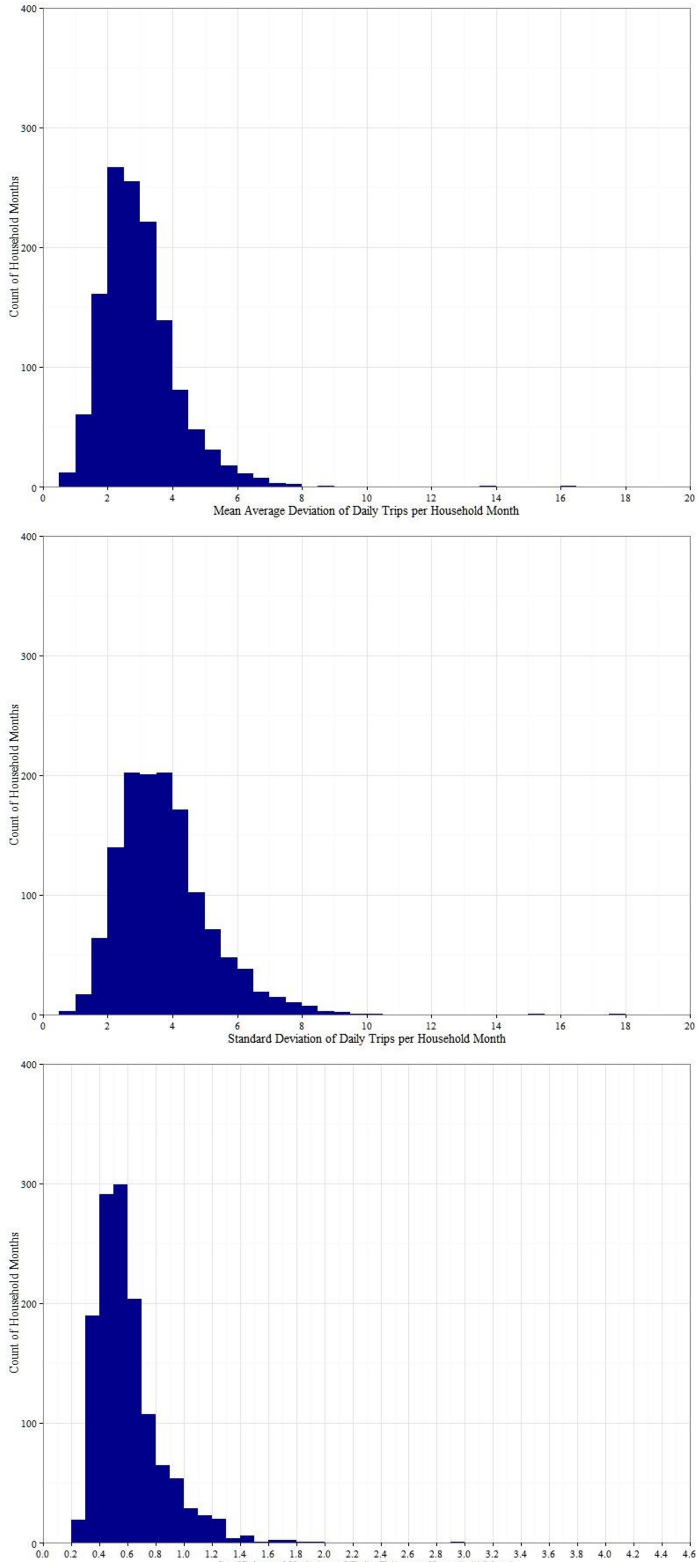


Figure 7.2 Distribution of Variability Measures of Daily Number of Trips per Household Month

Daily Vehicle Miles Traveled

The daily vehicle miles traveled provides the network length along which a household has traveled. While it is correlated to the number of activities it is more influenced by the extent of the activity locations. Figure 7.3 shows the distribution of daily vehicle miles of travel. The mean daily vehicle miles traveled is 49.2 miles and the mode of the distribution is approximately 30 miles. The mean daily vehicle miles traveled has a maximum value of 250 miles. One household had a consistently larger number of trips and daily vehicle miles of travel across many months of valid data. The household with large number of trips and daily vehicle miles traveled did not have any vehicles also used for commercial purposes.

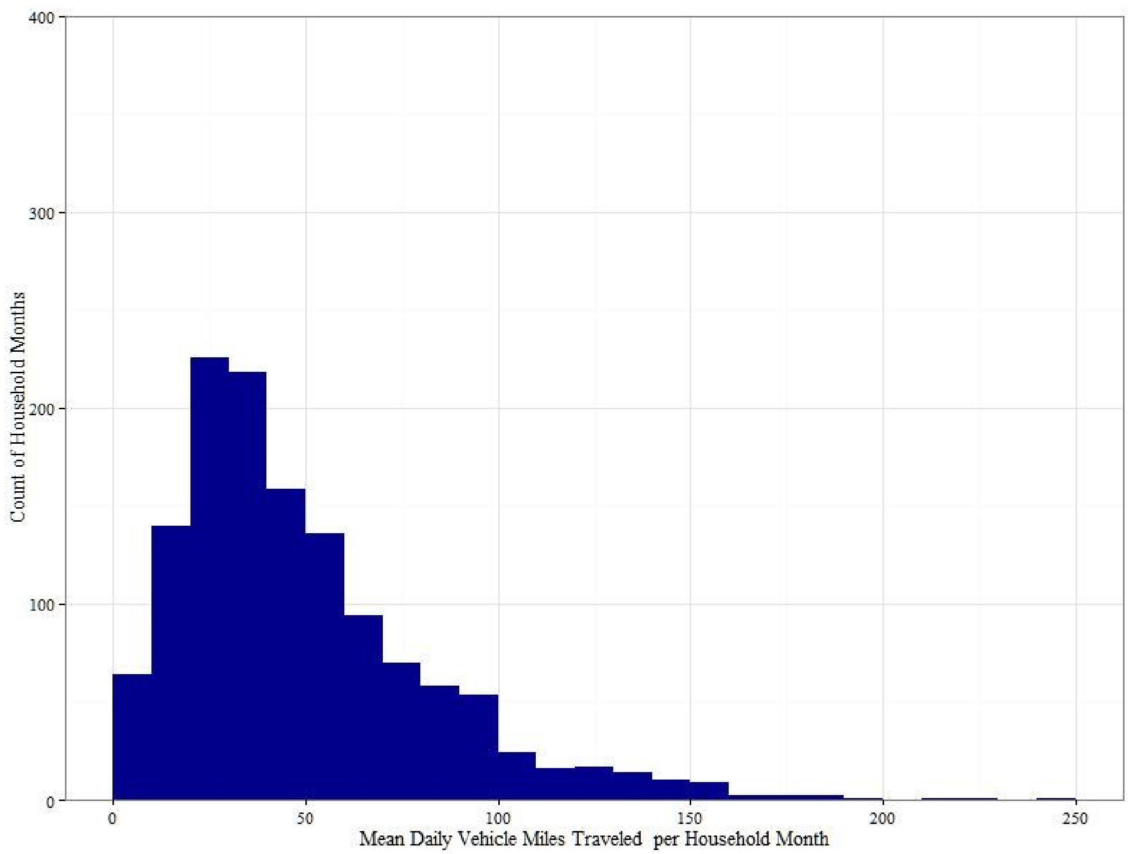


Figure 7.3 Distribution of Mean Daily Vehicle Miles Traveled per Household Month

Figure 7.4 presents the distribution of the mean absolute deviation, standard deviation and coefficient of variation of the daily vehicle miles traveled. Similar to the variability measures for the number of daily trips, the range of the mean absolute deviation of the daily vehicle miles traveled is smaller than the standard deviation of the daily vehicles miles traveled. The coefficient of variation of the daily vehicle miles traveled has a wider range than the coefficient of variation of the daily number of trips.

The mean of the mean absolute deviation of daily vehicle miles traveled is 27.8 miles and the mode is 20 miles. The mean of the standard deviation of daily vehicle miles traveled is 40.4 miles and the mode is 30 miles. The mean of the coefficient of variation of the daily vehicle miles traveled is 0.9 and the mode is 0.5.

The daily vehicle miles traveled may be correlated to the number of activities that a household participates. However it may not be as strongly correlated as the number of daily trips since it is also influenced by the spatial extent of activity locations. A household having fewer activities that occur farther apart can have more daily vehicle miles traveled than a household that has more number of activities that occur close to each other. Therefore, it may not be a better measure than number of daily trips to represent the travel behavior of a household.

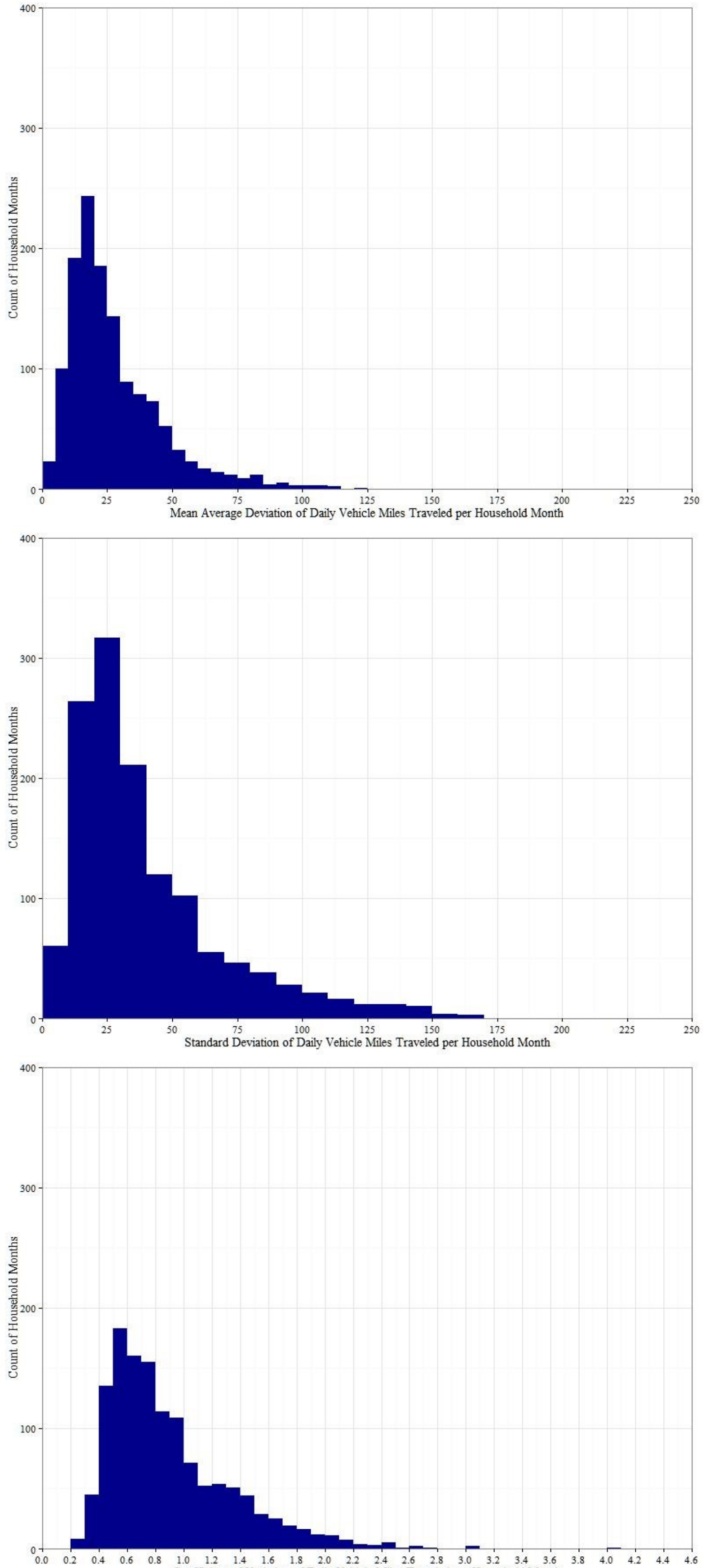


Figure 7.4 Distribution of Variability Measures of Daily Vehicle Miles Traveled per Household Month

Daily Vehicle Travel Time

The daily vehicle travel time provides the time spent on the road network every day. The daily vehicle travel time is influenced by the daily vehicle miles traveled, the characteristics of the road network and the time of day when the trip occurs. Figure 7.5 presents the distribution of the daily vehicle travel time for household month. The mean daily vehicle hours of travel by a household is two hours (121.2 minutes) and the mode is 1 hour and 20 minutes (80 minutes). The maximum mean daily vehicle travel by a household is 8 hours and 30 (508 minutes). While the maximum daily vehicle travel time appears large at first glance, the household demographics (high income, five vehicles, and six persons including two children) explains that household's behavior.

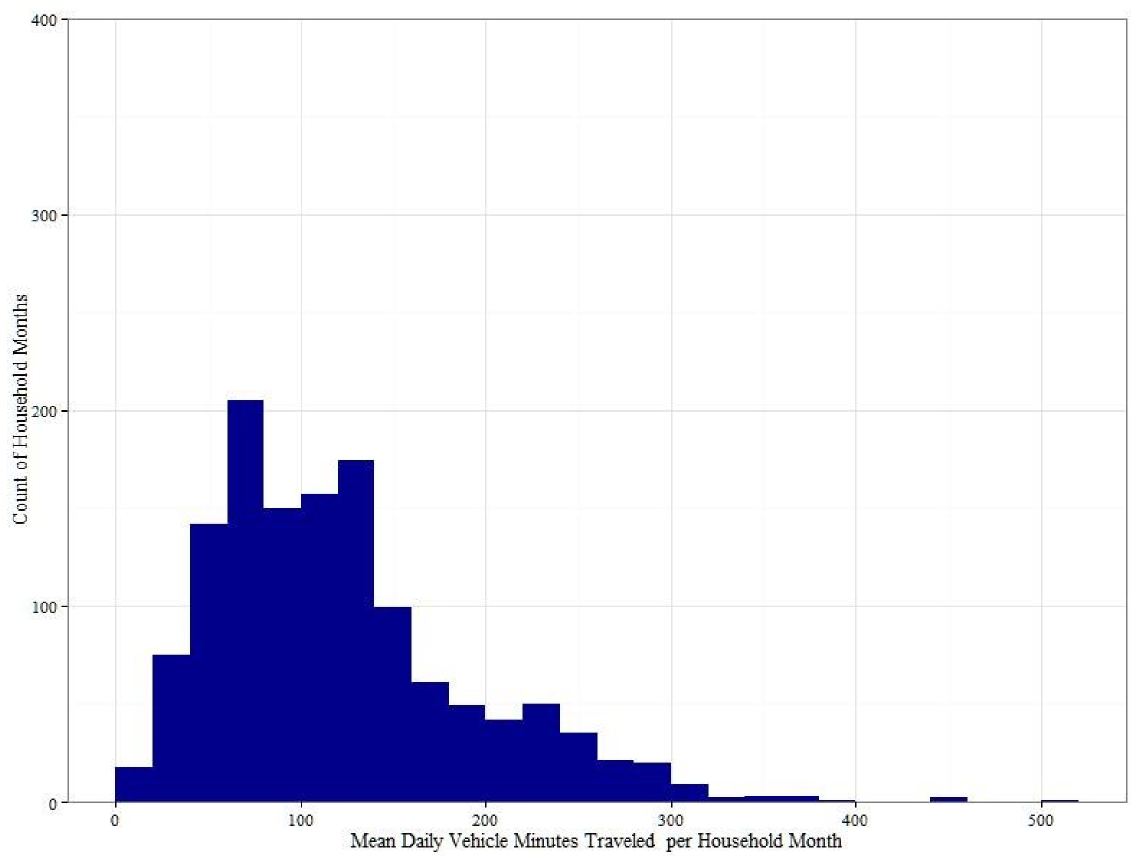


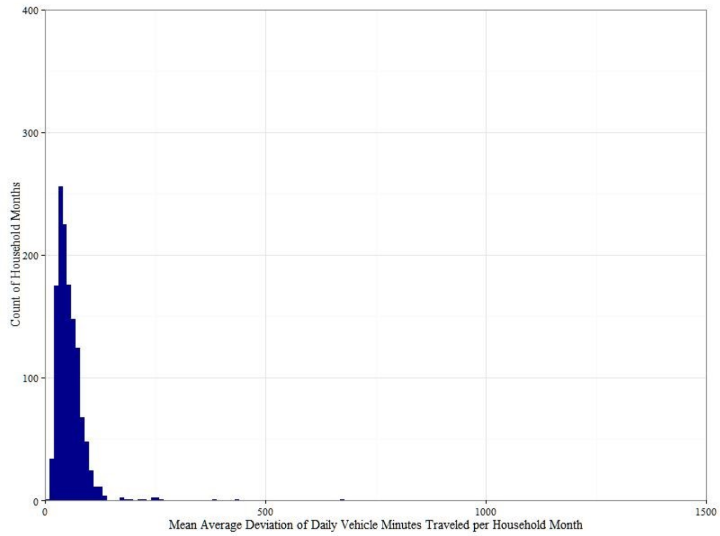
Figure 7.5 Distribution of Mean Daily Vehicle Hours of Travel (Expressed in Minutes) per Household Month

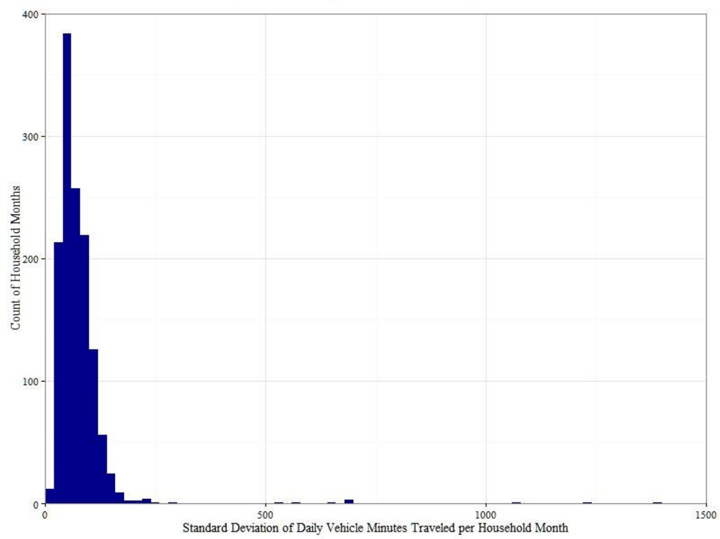
Figure 7.6 presents the distributions of the variability measures of the daily vehicle hours of travel (expressed in minutes). The range of the standard deviation of the daily vehicle hours traveled is larger than the mean absolute deviation of the daily vehicle hours traveled since standard deviation is influenced more by values farther from the mean than mean absolute deviation. The coefficient of variation of the daily vehicle hours traveled has a larger range than both the daily vehicle miles traveled and the daily number of trips.

The mean of the mean absolute deviation of the daily vehicle hours traveled (expressed in minutes) is 55.1 minutes and the mode is 60 minutes. The mean of the standard deviation of the daily vehicle hours of travel (expressed in minutes) is 75.3 minutes and the mode is 60 minutes. The mean of the coefficient of variation of the daily vehicle hours of travel is 0.68 and the mode is 0.60.

The daily vehicle hours of travel may be correlated with activities participation by the household. However, the daily travel time is significantly impacted by the roadway characteristics, congestion, and the distance traveled. A household having larger travel time may have most of the travel activity in congestion compared to a household that has more number of activities that occur under free flow speeds. Therefore, the daily travel time may not be a better measure to study activity participation than the number of daily trips or daily vehicle miles traveled.

The dissertation proposes to use the daily number of trips and daily vehicle miles traveled as travel behavior measures. The role of dispersion measures and demographics will be studied in detail with respect to number of trips per day since the daily vehicle miles traveled is correlated to the number of trips and can be expected to follow a similar pattern.





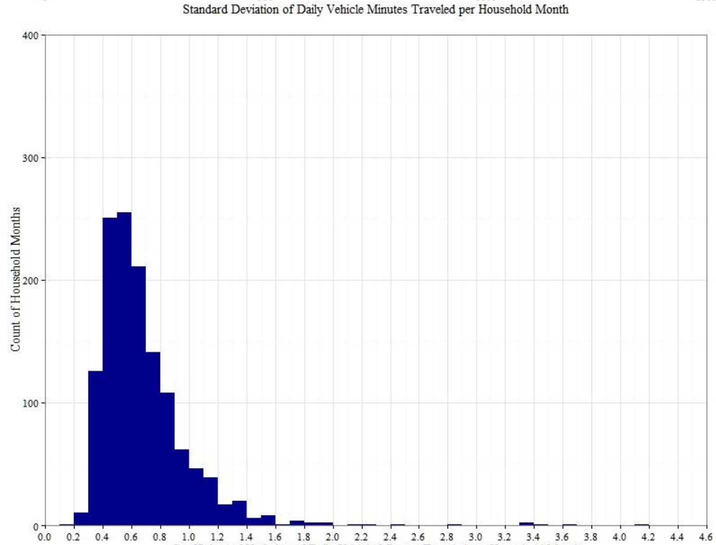


Figure 7.6 Distribution of Variability Measures of Daily Vehicle Hours of Travel (Expressed in Minutes) per Household Month

Evaluating Measures of Variability

This section explores the three variability measures of daily number of trips and evaluates the measures to capture the adventure seeking nature of the individual households. The variability measures for each household month come from the distribution of the daily travel characteristics of each household month. The mean of the distribution of number of trips differs across the various household months. Therefore it is important to study the relationship between the mean of number of trips distribution for each household month and the variability measure of the number of trips distribution.

Mean Absolute Deviation of Number of Daily Trips

Figure 7.7 shows the scatter plot of Mean Absolute Deviation of daily trips per household on the Y-axis and the mean daily trips per household on the X-axis. From figure 7.7 we can see that the slope of the regression line is 0.24 which implies that the mean absolute deviation increases linearly with the mean number of daily trips for a household. Therefore, it can be concluded that the mean absolute deviation may be influenced by the mean daily number of trips for a household. Two households with largely different mean daily number of trips but the same mean absolute deviation do not have the same amount of variability in their travel behavior. The R^2 value of 0.59 indicates a reasonable fit between the regression line and the scatter plot.

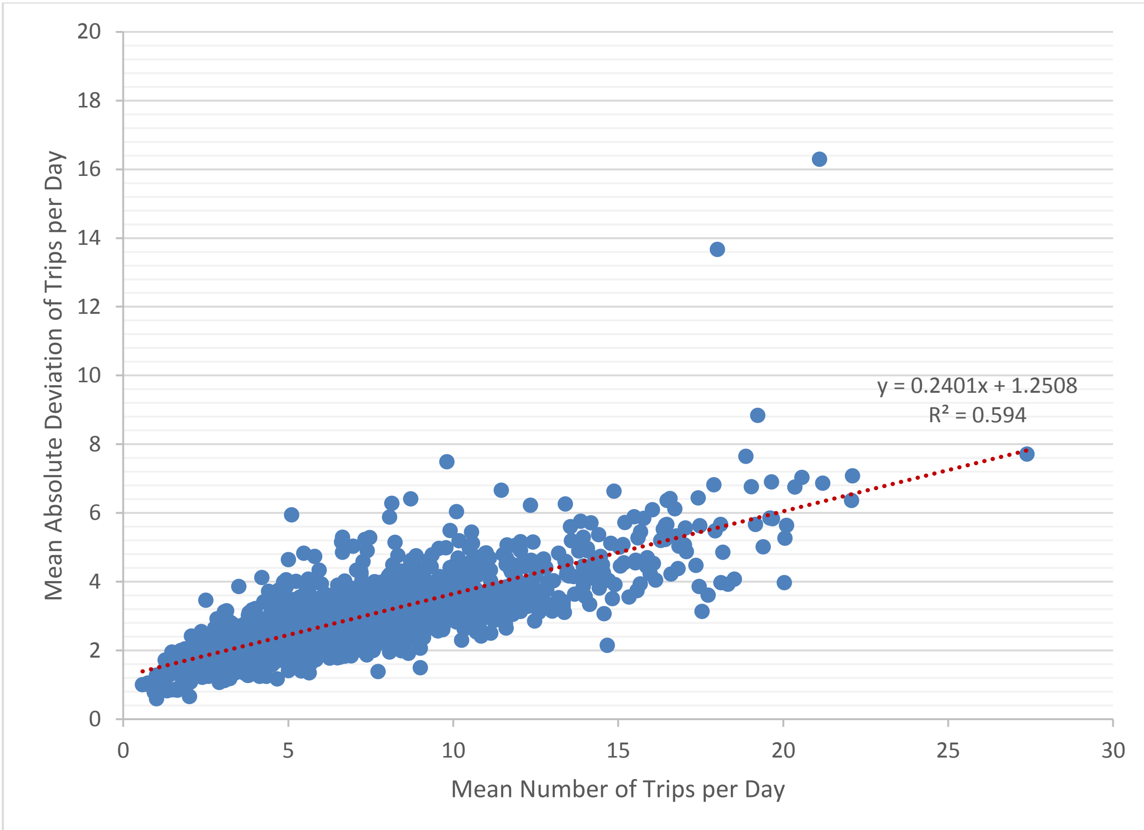


Figure 7.7 Scatter plot of Mean Daily Number of Trips and Mean Absolute Deviation of Daily Number of Trips

Standard Deviation of Number of Daily Trips

Figure 7.8 shows the scatter plot of standard deviation of the daily number of trips and the mean daily number of trips. The slope of the regression line is 0.3 indicating the possible influence of the mean number of daily trips on the standard deviation. The R^2 value is 0.62 for the regression line which is slightly higher than the R^2 value of the regression line in the scatter plot between mean absolute deviation and mean daily number of trips.

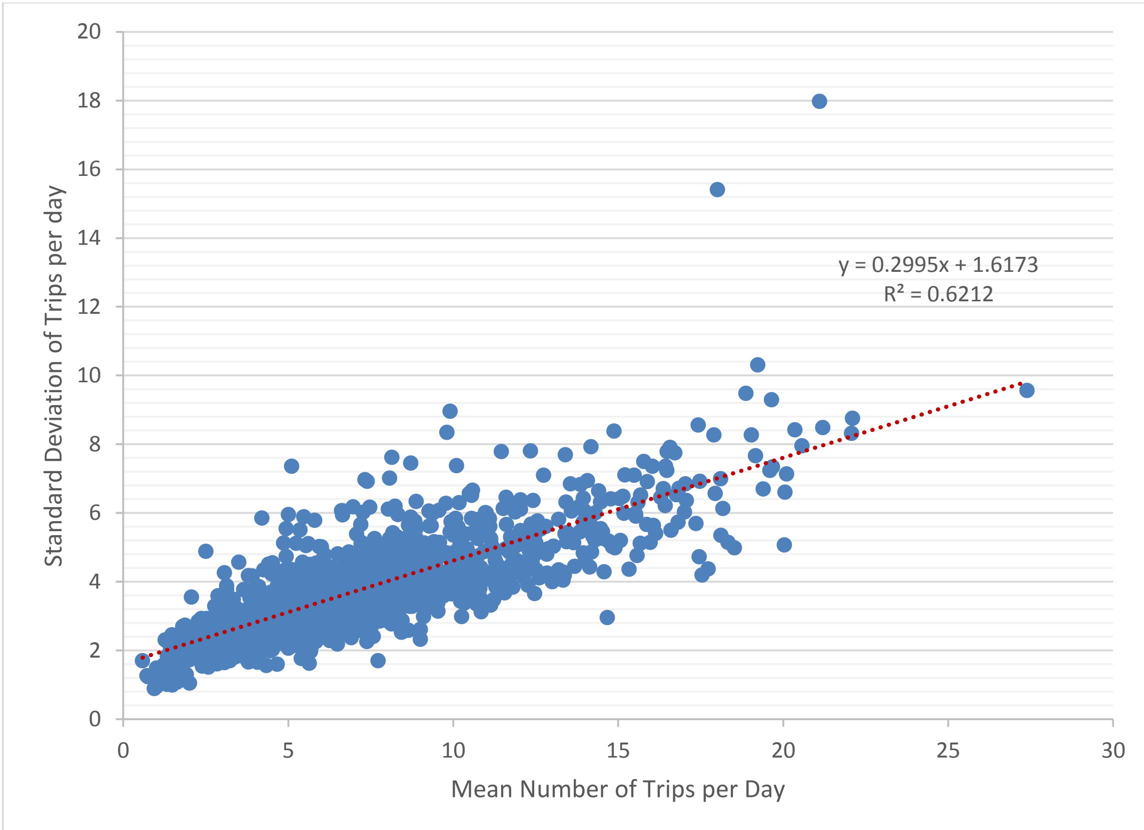


Figure 7.8 Scatter plot of Mean Daily Number of Trips and Standard Deviation of Daily Number of Trips

Coefficient of Variation of Number of Daily Trips

Figure 7.9 shows the scatter plot between coefficient of variation of the daily number of trips and the mean number of trips. The coefficient of variation of trips per day is larger when the mean number of daily trips is less than five trips/day and as the mean number of daily trips increases the coefficient of variation of trips per day decreases and ultimately stabilizes. The linear regression line has a slope of -0.04 and the R^2 value is 0.40 which suggests a lesser impact of the mean daily trips on coefficient of variation of trips per day compared to the mean's impact on the mean absolute deviation or the standard deviation of the daily number of trips. Based on the scatterplot between the coefficient of variation of the trips and the mean number of daily trips, a power regression line was tested which resulted in a R^2 value is 0.60. The power regression equation represents an asymptotic curve. The coefficient of variation is likely influenced by the mean number of trips, especially when the mean number of trips per day is less than 5. The number of trips is a discrete value and even a variation of one trip by households that make few trips can lead to a significantly large coefficient of variation. Hence, this method may lead to some bias for households that make few trips.

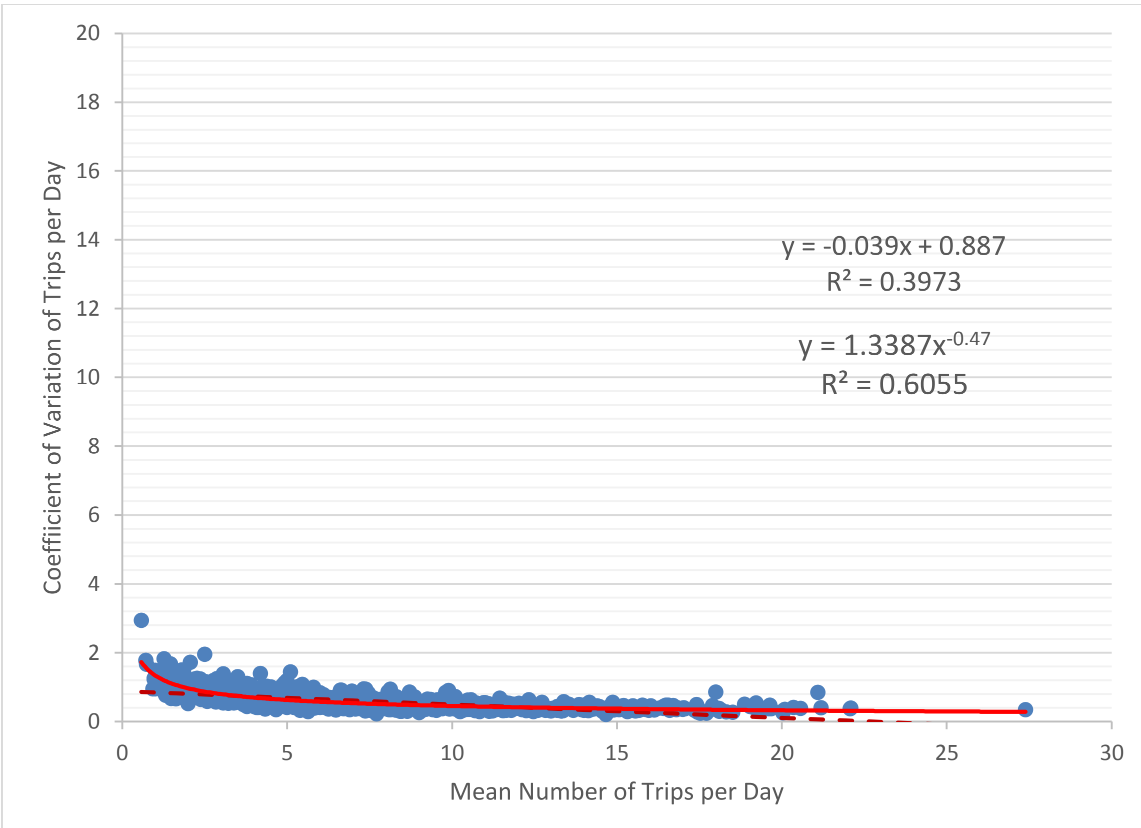


Figure 7.9 Scatter plot of Mean Daily Number of Trips and Coefficient of Variation of Daily Number of Trips

The Spearman's Rho, a non-parametric correlation test was applied to test the correlation between the mean daily number of trips and measures of variability. The results are tabulated in table 7.1. The mean absolute deviation and standard deviation show a positive correlation with the mean number of daily trips and the coefficient of variation of trips per day has a negative correlation with the mean number of daily trips. The absolute correlation value is highest for the standard deviation and lowest for the coefficient of variation.

Table 7.1 Correlation Tests between Mean Number of Daily Trips and Measures of Variability

Measure of Variability	Spearman's Rho	P-value
Mean Absolute Deviation	0.78	<0.001
Standard Deviation	0.80	<0.001
Coefficient of Variation	-0.76	<0.001

From the above analyses, all the three measures of variability have some correlation with the mean number of daily trips. However, the standard deviation appears to be most influenced, among the measures of variability, by the mean number of daily trips based on the slope of the regression line and the Spearman's Rho. The coefficient of variation of trips per day may be the least influenced measure of variability with a very small slope of -0.04 and the least absolute value of the Spearman's Rho. Figure 7.8 also shows that the coefficient of variation of trips per day decreases and stabilizes once the mean number of daily trips increases beyond five trips/day. This infers high travel variability for households that make fewer trips, which is intuitive. The coefficient of variation of trips per day appears to be the most suitable measure to study the variability in household travel behavior. However, quite a few studies have used variance or standard deviation as the measure of variability [4, 13, 20, 24, 63] and it may be worth the effort to evaluate using that measure in activity participation modeling.

Travel Behavior Variability and Demographics

Prior research efforts have shown that the travel behavior variability is influenced by household socio-economic and demographic characteristics [4, 13, 20, 22, 24, 48, 59]. It is important to understand the potential relationships between the demographic attributes and the measures of variability before using the measure of variability as a surrogate for the variability-seeking nature of a household in the modeling process. The relationships will help in classifying the measure of variability as an exogenous or endogenous variable in the modeling process. The previous section found that the coefficient of variation of trips per day may be the most suitable measure of variability. Hence, this section will explore its relationships to the demographic and socio-economic characteristics.

Household Income

Household income is directly proportional to the number of trips. Households with higher income generally undertake more discretionary activities. Figure 7.10 shows the monthly mean coefficient of variation of trips per day by household income with the bootstrap confidence bounds. The coefficient of variation of trips per day is higher for lower income households since those households tend to have lower number of daily trips which results in higher coefficient of variation of trips per day as seen in previous section. The confidence bounds for the lower income group does not overlap with the confidence bounds of the higher income group and thus represents statistically different coefficient of variation of trips per day between the two groups. The confidence bounds are not uniform above and below the mean since the distribution is not normally distributed and this asymmetry is effectively captured by bootstrap methods.

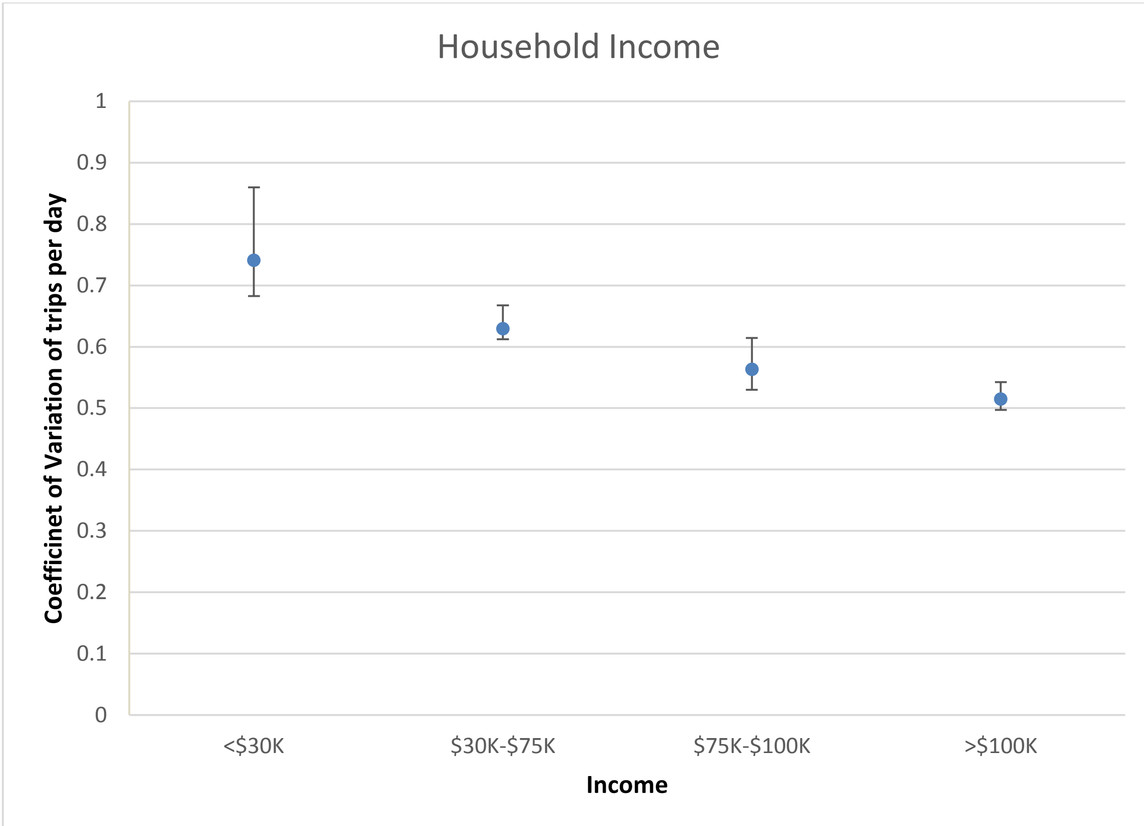


Figure 7.10 Bootstrapped Confidence Intervals for the Mean Coefficient of Variation of Trips per Day by Household Income

The results of the Mann Whitney U test to test the null hypothesis that the coefficient of variation of trips per day is similar between different income groups are shown in Table 7.2. The null hypothesis is rejected at the 95 percent confidence level across all income group comparisons. These results indicate that the coefficient of variation of trips per day is influenced by the household income.

Table 7.2 Mann Whitney Test Comparing Coefficient of Variation of Trips per Day by Household Income

Group 1	Group 2	N1	N2	Rank Sum 1	Rank Sum 2	P-value
<$30K	$30K-$75K	131	706	62663	288040	0.002
$30K-$75K	$75K-$100K	706	182	326816	67899	0.000
$75K-$100K	$100K+	182	300	47013	69389	0.039

Household Size

The household size is proportional to the number of daily trips. The presence of more household members generally necessitates more activities by the household. Figure 7.11 shows the mean coefficient of variation of trips per day by household size with the confidence bounds of the mean calculated by bootstrap methods. The one person household has the highest mean coefficient of variation of trips per day and households with three or more people tend to have similar coefficient of variation of trips per day.

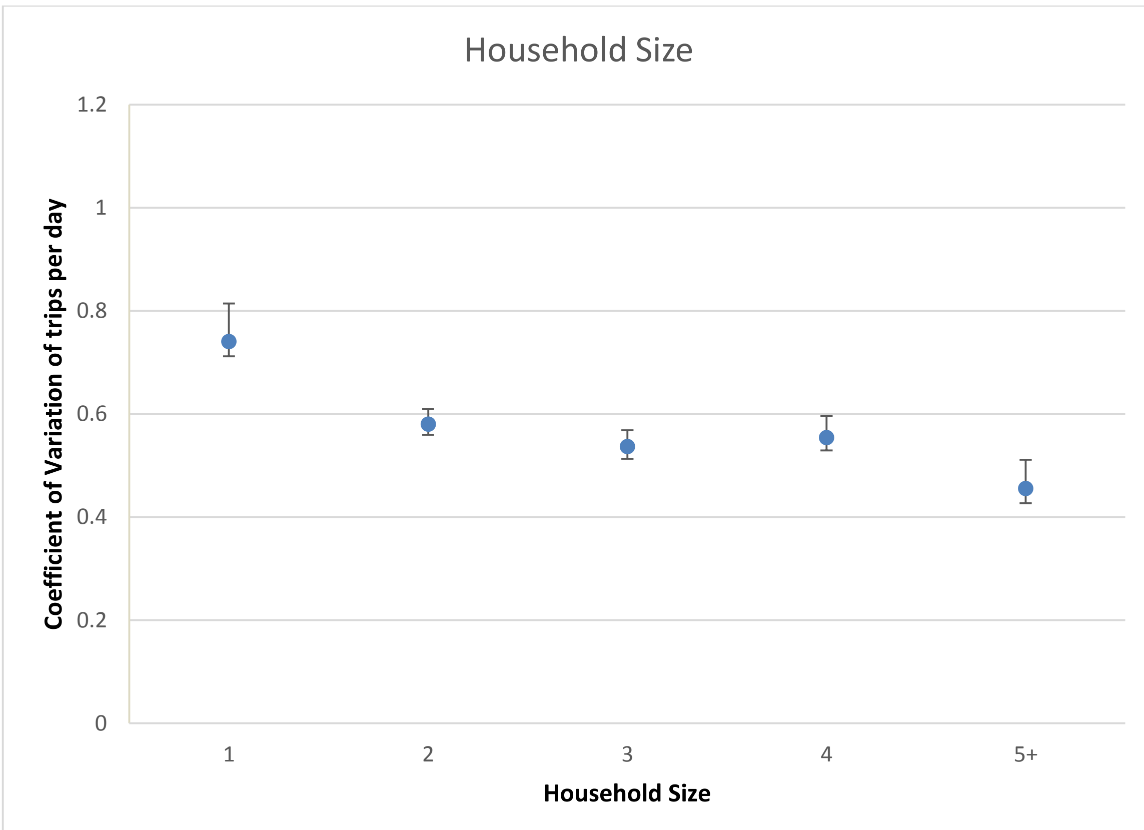


Figure 7.11 Bootstrapped Confidence Intervals for the Mean Coefficient of variation of Trips per Day by Household Size

Table 7.3 presents the results of the Mann Whitney comparing the coefficient of variation of trips per day across the different household size groups. The null hypotheses that the coefficients of variation are similar across the different household size groups are tested. The results indicate that the coefficient of variation of trips per day is different across the household sizes, except for households with three and four persons. Households with three and four persons may have more similar characteristics in terms of household structure and lifecycle stages, due to the presence of children. These results indicate that household size may affect the coefficient of variation of trips per day.

Table 7.3 Mann Whitney Test comparing Coefficient of Variation of Trips per Day by Household Size

Group 1	Group 2	N1	N2	Rank Sum 1	Rank Sum 2	Probability
1 person	2 persons	369	455	185191	154709	0.000
2 persons	3 persons	455	200	154446	60394	0.020
3 persons	4 persons	200	205	39912	42303	0.559
4 persons	5+ persons	205	90	33669	9991	0.000

Household Vehicle Ownership

Household vehicle ownership is another important variable that is proportional to the number of trips made by a household. Household vehicle ownership is also generally correlated to the household income and household size. Figure 7.12 shows the mean coefficient of variation of trips per day by household vehicle ownership. The confidence bounds around the mean are calculated using bootstrap methods. With increase in the vehicle ownership the coefficient of variation of trips per day decreases similar to household income and household size.

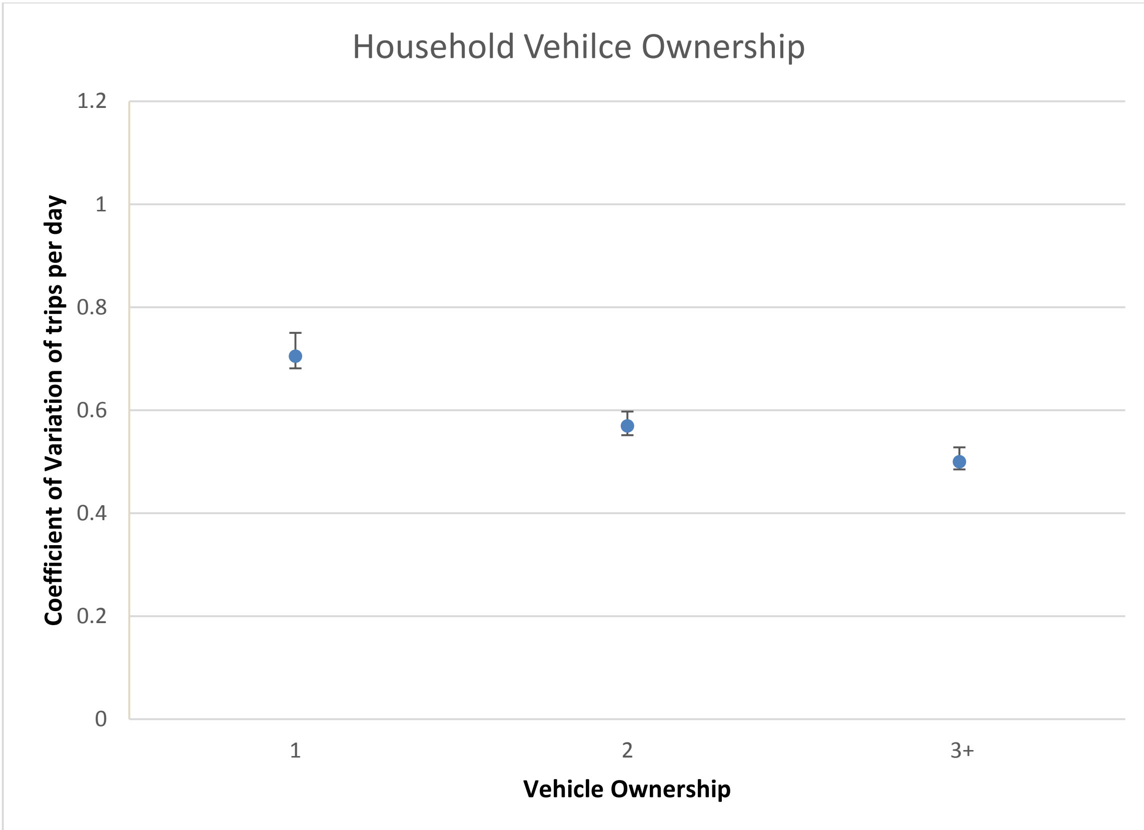


Figure 7.12 Bootstrapped Confidence Intervals for the Mean Coefficient of Variation of Trips per Day by Household Vehicle Ownership

The results of the Mann Whitney test to compare the coefficient of variation of trips per day by household vehicle ownership are shown in Table 7.4. The null hypothesis that the coefficient of variation of trips per day is the same across different household vehicle ownership groups is rejected at the 95 percent confidence level. The household vehicle ownership may affect the coefficient of variation of trips per day.

Table 7.4 Mann Whitney Test Comparing Coefficient of Variation of Trips per Day by Household Vehicle Ownership

Group 1	Group 2	N1	N2	Rank Sum 1	Rank Sum 2	Probability
1 vehicle	2 vehicles	506	497	298408	205097	0
2 vehicles	3+ vehicles	497	316	218120	112770	0

Number of Children

The presence of children in a household influences the activities that the household undertakes and therefore affects the number of daily trips for that household. The presence of children is also correlated with household size. Figure 7.13 presents the results of the mean coefficient of variation of trips per day by the presence of children. The confidence bounds around the means were calculated by bootstrap methods. Households without children exhibit significantly higher daily travel variability than households with children even though the average number of trips with children is 8.46 trips per day and without children 6.57 trips per day. Households with children have more trips per day but the travel behavior is more routine and therefore they have less variability. On the other hand household without children pursue more activities that are not routine and can satiate their desire to seek variability in travel.

The null hypothesis that the coefficient of variation of trips per day is the same across households with and without children was assessed using the Mann Whitney U test and the results are shown in Table 7.5. The null hypothesis was rejected at the 95 percent confidence level, suggesting that the presence of children may affect the variability in travel behavior.

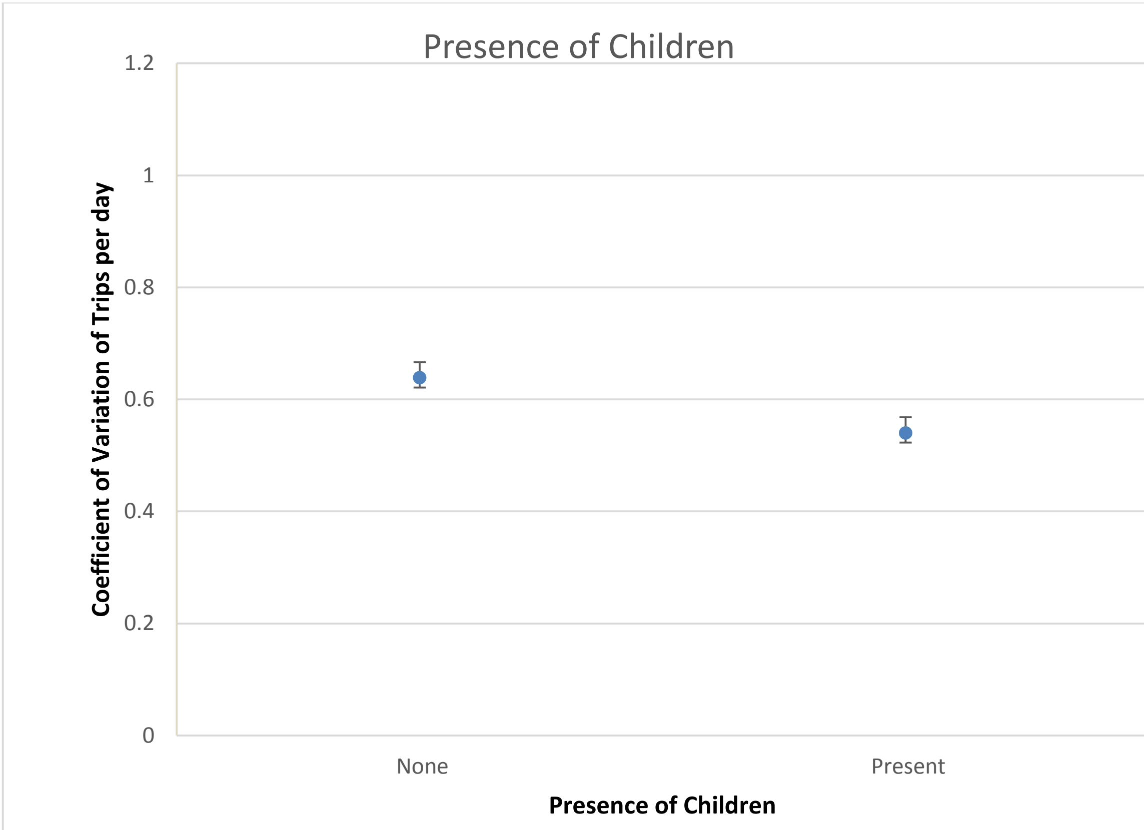


Figure 7.13 Bootstrapped Confidence Intervals for the Mean Coefficient of Variation of Trips per Day by Presence of Children

Table 7.5 Mann Whitney Test Comparing Coefficient of Variation of Trips per Day by Presence of Children

Group 1	Group 2	N1	N2	Rank Sum 1	Rank Sum 2	Probability
Children Present	No Children	447	872	247050	623489	0.000

Workers

The presence of workers in households may indicate the conduct of more habitual travel behavior, compared to households that do not have workers. With more habitual behavior, variability in travel behavior is expected to decrease. Figure 7.14 shows the mean coefficient of variation of trips per day by the presence and absence of workers in the households, with the confidence bounds around the means calculated by bootstrap methods. Households with no workers have higher variability in travel than households that have at least one worker.

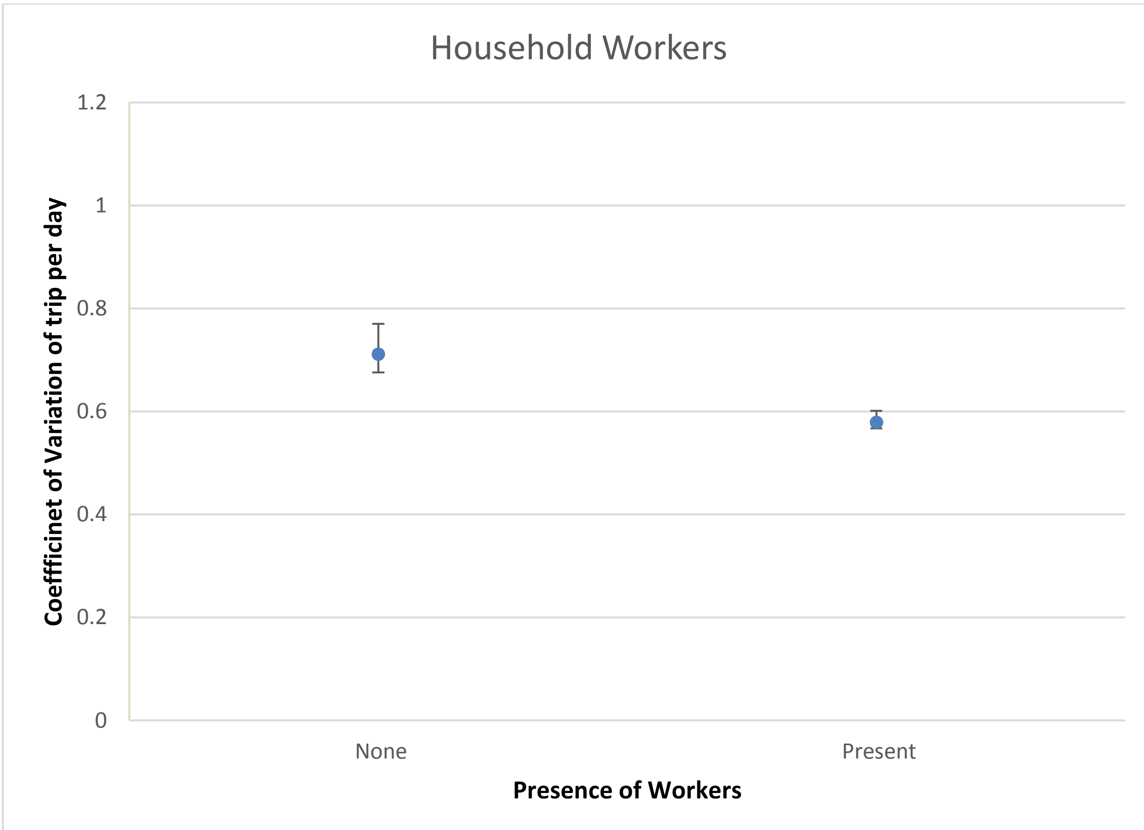


Figure 7.14 Bootstrapped Confidence Intervals for the Mean Coefficient of Variation of Trips per Day by Presence of Workers

The results of the Mann Whitney test to compare the coefficient of variation of trips per day by the presence of workers are shown in table 7.6. The null hypothesis that the coefficient of variation of trips per day are similar between the two groups is rejected at the 95 percent confidence level. The presence or absence of workers in households appears to affect the household travel behavior variability.

The household income, household size, household vehicle ownership, presence of children, and presence of workers all affect the coefficient of variation of trips per day. This implies that the coefficient of variation of trips per day needs to be used as an endogenous variable in the modeling process.

Table 7.6 Mann Whitney Test Comparing Coefficient of Variation of Trips per Day by Presence of Workers

Group 1	Group 2	N1	N2	Rank Sum 1	Rank Sum 2	Probability
Workers Present	No Workers	1054	265	647027	223513	0.000

Summary

The first section in this chapter presented the results of the variability measures of the travel behavior attributes, including: number of daily trips, daily vehicle miles traveled, and the daily vehicle hours traveled. The analytical results indicate that daily vehicle miles traveled may be influenced by the distance between the activity locations, and not just the number of activities. The daily vehicle hours traveled may be affected by the road characteristics, congestion and distance traveled. The number of trips per day and the daily vehicle miles traveled may be the most suitable travel behavior measure to capture the overall travel behavior of an individual household.

The second section evaluated the mean absolute deviation, standard deviation and the coefficient of variation of the number of daily trips. The mean absolute deviation and the standard deviation had a significant slope on the scatter plot with the number of trips per day. Two households with largely different calculated mean daily number of trips but the same mean absolute deviation or standard deviation do not exhibit the same amount of variability in their travel behavior. Hence, using the absolute mean deviation or mean standard deviation may not clearly identify differences in variability across such households. In the analyses conducted with data from these 95 households, the coefficient of variation of trips per day increases as the number of trips drops below five trips per day. A power regression line best fits the relationship between coefficient of variation trips per day and the number of trips per day which indicates that the curve is asymptotic. The coefficient of variation of trips per day and the coefficient of variation of daily vehicle miles traveled may be the most suitable measure to compare travel behavior variability between households.

The third section evaluated the coefficient of variation of trips per day and its relationship with household demographics and socio-economic characteristics in detail. Household income, household size, and household vehicle ownership are all correlated with the coefficient of variation of trips per day. The presence or absence of children and workers were also found to impact the coefficient of variation of trips per days. These results suggest that the coefficient of variation of trips per day needs should be considered as an endogenous variable in the modeling process.

CHAPTER 8

METHODOLOGY TO ESTIMATE ACTIVITY SPACE

The next step in this dissertation is to estimate the spatial extent of activity for the participating households to use it in the activity participation model building process. Advances in travel behavior analysis and statistical methods have set the focus on disaggregate analysis of travel survey data [59]. Most disaggregate travel behavior studies have measured behavior in terms of activity types, number of trips, and distance traveled. However, with increased computing power and availability of geographical travel data from GPS based travel surveys, spatial travel behavior analysis is coming to the forefront. Spatial analysis answers the 'Where' part of the 'When', 'Where', 'How' and 'Why' questions that planners and modelers are trying to answer about travel behavior. While spatial activity extent is influenced by socio-economic characteristics, it is also affected by environmental factors, such as location of the household, and the inherent spatial appetite of household members. Estimating the activity space extent and using it in the model building process can help us better explain activity participation in these models.

This chapter develops a methodology to integrate high resolution GPS travel data into activity space estimation. The first section explains the concept of activity space in terms of travel behavior. The second section details the two most commonly used methods in travel behavior activity estimation, namely the Confidence Ellipse and Kernel Density methods. The next section outlines the assumptions made in estimating the activity space for the Commute Atlanta data set. The next section explores the application of the two methods of activity space estimation to the Commute Atlanta data set and discusses the limitations of those methods. The same section also proposes a new methodology to estimate activity space called "Modified Kernel Density Method" that is designed to address the limitations in the other two methods. The last section summarizes the research efforts in this chapter.

Activity Space

Activity space is the geographic area within which the individuals in a household live and interact on a daily basis. The size of the activity space is affected by socio-economic characteristics, characteristics of the built environment, season, and the individuals' choice of activity locations. Mobility patterns from traditional two-day travel diary surveys do not capture all habitual activity which typically have longer cycles. With increased use of longitudinal travel surveys, it is possible to measure the activity survey and study the usage of urban space [72].

Traditionally travel behavior activity space is estimated using the spatial distribution of activity locations [72]. However, the driving activity itself is not included in these methods due to the lack of detailed route data from most travel surveys. The built environment around the driving activity may influence future activities (i.e. trip chaining) and may influence the route choices people make [73]. Activity space estimated by traditional methods does not directly capture the potential impacts of the actual travel as integral to the activity. However, technological advances have now made high resolution GPS data widely available to researchers.

Activity Space Estimation Methods

As described in Chapter 2, activity space is often measured as the area under the Confidence Ellipse, the area of the cells with a minimum density in the Kernel Density method, or by the length of the minimum spanning trees [72]. This dissertation effort primarily focuses on the Confidence Ellipse and the Kernel Density methods.

Confidence Ellipse

Confidence Ellipses are analogous to the confidence interval of univariate distributions as the smallest possible (sub-)area in which the true value of the population should be found with a certain probability (e.g. 95%) [28, 74]. The measure of activity space is the area within the Confidence Ellipse. Confidence Ellipse is a measure of dispersion and is also called as standard deviation ellipse [75]. The Confidence Ellipse method assumes normally distributed spatial data and is the confidence region of two dimensional bivariate normal distribution [74]. The mathematical structure of Confidence Ellipse estimation is detailed in Chapter 2 as part of the literature review.

The area under the Confidence Ellipse also includes large areas that the individual is not aware of and into which the individual will never venture. Areas that are not connected through the network, and areas between activity locations, add to the size of the Confidence Ellipse. The ability of Confidence Ellipse to effectively represent the travel behavior is limited by the addition of these extra areas. Confidence ellipses may be used to compare the dispersion of activity locations between individuals or compare the activity space of an individual between different periods of time [74]. However, Confidence Ellipse does not reflect where the individual actually visits.

Rai, et al., examined the use of various new geometries in capturing human activity space [76]. The authors in that research effort looked at using the ellipse, super-ellipse, Cassini oval, and the Bean curve. The study tested the four parametric geometries which captured 95% of the locations visited while minimizing the area under the geometry. The research concluded that none of the geometries fits for all households, and the ellipse is not a good representation of the spatial activity.

Kernel Density

The Kernel Density method is a non-parametric method that does not assume any distribution for the underlying locations [74]. Kernel densities are the transformation of a point pattern, such as a set of activity locations, into a continuous representation of density in a wider area [28]. The Kernel Density area is the sum of all areas with at least certain non-zero probability of activity occurrence. The estimation of Kernel Density is a smoothing technique, which generalizes observation points to the area in which they are found.

There are multiple approaches to the estimation of Kernel Density. One popular method is the Fixed Kernel Method [77]. In this method, a variably distributed kernel function is placed over each data point. The sum of the overlapping values gives the density estimate. The mathematical layout of this method is detailed in Chapter 2.

The Kernel Density gives a good distribution of the frequency of the activity locations and the activity area. The Kernel Density method captures the area surrounding the activity locations and unlike the Confidence Ellipse method does not include area where the individual has never visited. However, the Kernel Density does not effectively capture the dispersion of the activity locations.

Assumptions for the Commute Atlanta Data Set

The activity space estimated is expected to encompass all habitual activity locations of an individual household. If all habitual activities are not captured the activity space may not represent the complete spatial appetite of the household and will affect the results of models that use the activity space estimates. The following assumptions will be made for developing a methodology to estimate activity space.

- For the purpose of this research effort, all household trips are motor vehicle trips. The Commute Atlanta instrumented vehicle data collection effort did not collect walking, cycling, or transit travel data. As discussed in previous chapters the primary mode of transportation in Atlanta is the personal vehicle and households that own vehicles have a higher share of personal vehicle trips than the overall population. The activity space generated by this study will miss activities and travel undertaken by other modes.
- One month of travel data will encompass all habitual activity patterns. Schlich, et al., found that at least two weeks of travel data to encompass most habitual activity patterns [14] and Schoenfelder et al. found that the number of new locations visited stabilized at 22 days in Atlanta [3]. In the Commute Atlanta study 70% of the 95 households experienced changes to socio-economic characteristics, vehicle ownership, and location changes during the study period [48]. Hence it is also necessary to limit the longitudinal bins to one month to simultaneously incorporate the effects of the demographic and location changes over time [78].
- The maximum walking distance from the end of vehicle trip to activity location is one quarter mile (approximately 400 m). Atash reports 0.25 miles or 400 m has been assumed and accepted as the distance that an average American will walk rather than drive and this distance has been used by other studies to estimate pedestrian accessibility [79-81]. It is reasonable to expect individuals to interact within a quarter mile radius of the vehicle trip end.
- The maximum distance that an individual can visually interact by seeing from inside a vehicle during a trip is 300 feet (approximately 100 m). This is based on the assumption that individuals can effectively see commercial facilities and points of interest that are along

the road network and they may choose to pursue an activity in the future based on the visual information that they obtained during the travel activity.

Methodology to Estimate Activity Space in Commute Atlanta Study

The activity spaces for this research effort where estimated for all spatial locations visited each calendar month. The Commute Atlanta Study collected data over three years and about 70 percent of the households had demographic or vehicle changes during the study period. It is expected that a period of one month is long enough to capture most habitual activity cycles and fine enough to separate out the effects of changes in demographics and seasonal impacts s activity space.

ArcGIS was used to estimate the activity space by Confidence Ellipse and Kernel Density methods. The Confidence Ellipse or standard deviational ellipse was calculated using the "DirectionalDistribution_stats" function of ArcGIS software. The area that covered two standard deviations or 95% of the activity locations was the output of the function. The frequency of each location visited was used to weight their significance in the Confidence Ellipse estimation. Figure 8.1 shows a Confidence Ellipse generated for a sample household with all activity locations visited in March 2005. The confidence ellipse covers 393.7 sq Km with a standard distance of 1.5 Km along the minor axis and standard length of 8.3 Km along the major axis.

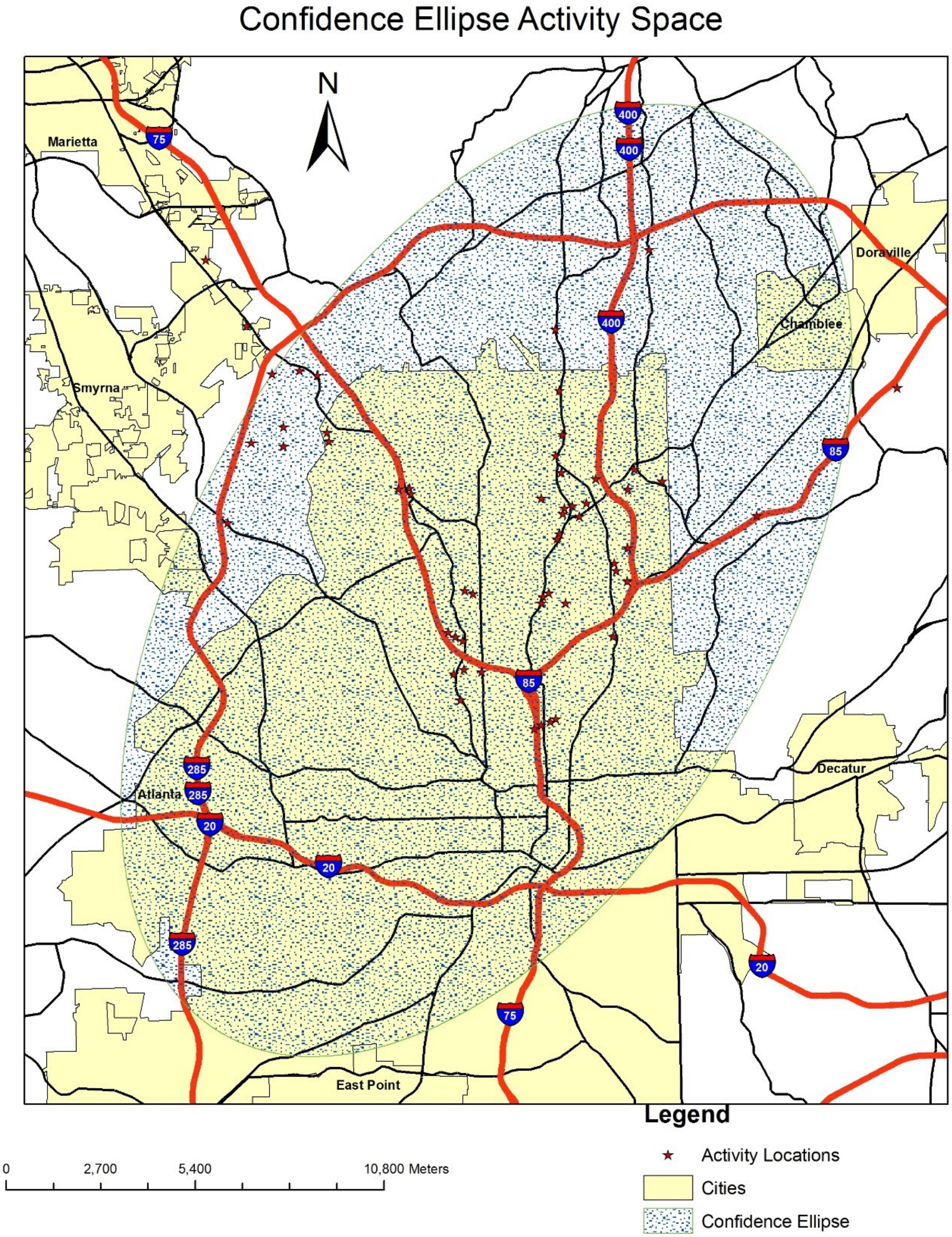


Figure 8.1 Confidence Ellipse Generated for Sample Household

The Kernel Density was calculated over a grid of 100m by 100m cells with a 'search radius' or bandwidth of 400m. The search radius of 400m is approximately a quarter mile following the assumption that an individual is likely to walk within a quarter mile from where they park their vehicle [79-81]. The frequencies of activity at each of the trip end locations were used as weight in the density calculations. To estimate the Kernel Density area, all cells that have non-zero values were included. Figure 8.2 shows the Kernel Density generated for the sample household which was used to represent the Confidence Ellipse area in Figure 8.1. The Kernel Density area for this sample household is 21.9 sq Km.

The purpose of the Confidence Ellipse method is to measure the dispersion of the activity locations whereas the purpose of the Kernel Density method is to measure clusters of activity locations and the activity space associated with those clusters. Both Confidence Ellipse and Kernel Density methods are weighted by the frequency of the activities at each location. Confidence Ellipse areas are of much larger magnitude since the Confidence Ellipse includes areas never visited by the individual. Kernel Density area is much more focused on the activity locations and present the clustering of activity attractions for a household.

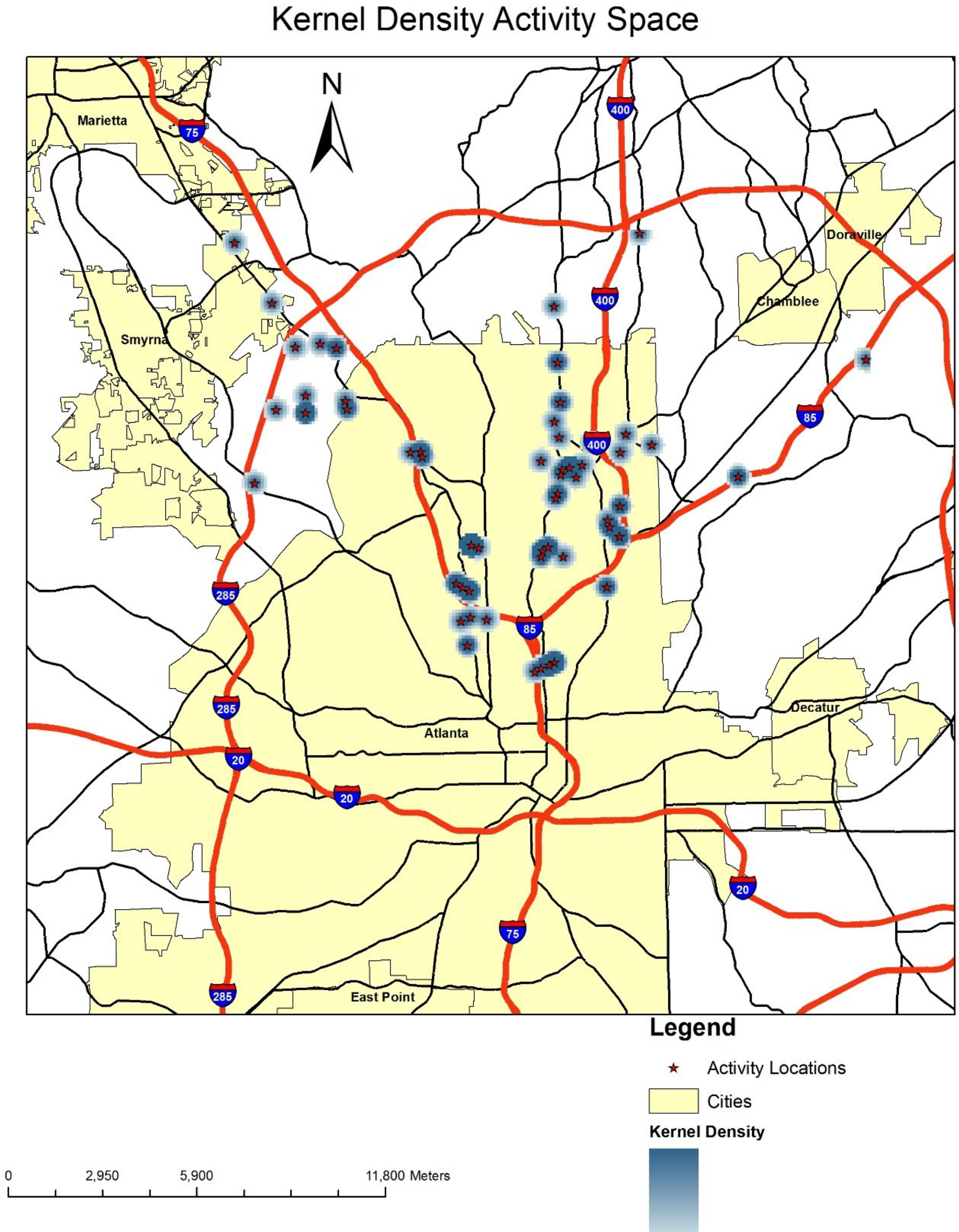


Figure 8.2 Kernel Density Area Generated for Sample Household

Modified Kernel Density Area

The area under the Confidence Ellipse reflects the spatial extent of the activity locations and does not effectively differentiate the activity frequency between two households that have the same spatial extent. Even when weighted by the frequency of activity at each location, it is possible to have two households with very different numbers of trips and therefore activities to yield the same Confidence Ellipse area. The Confidence Ellipse also covers a lot of area with which the individual never interacts. The Kernel Density area reflects the activity frequency and the number of activity locations. However, it does not capture the spatial extent of these activities, especially when using a small search radius and when the activities are widespread.

The research effort proposes a new methodology, which will address the spatial extent and frequency of activities with a single measure by considering the travel between activities as an activity. This study proposes a new measure, Modified Kernel Density, which includes the activity locations at trips ends and the space through which the travel activity is conducted.

The Kernel Density area for the trip end activity locations were estimated using the quarter mile radius and the Kernel Density area for the travel activity were estimated using a 300 feet radius which is the distance that is visible to the participant while engaging in the driving activity. The Comprehensive Kernel Density Area can be represented by

$$Modified\ KD = KD_{act} + KD_{travel}$$

Where,

$Modified\ KD$ Modified Kernel Density Area

KD_{act} Kernel Density Area of activity locations at Trip end using quarter mile radius

KD_{travel} Kernel Density Area around the travel activity using 300 feet radius.

The use of different radius in the Kernel Density area estimation is to reflect the nature of interaction with the space during the activities. At the end of the trips, the individual interacts with the space and is assumed to be willing to travel as far as a quarter mile on foot. During the travel activity, the individual is assumed to be limited to visual interaction with the space around him and is unlikely to see anything beyond the immediate environment along the road. Figure 8.3 shows the Modified Kernel Density Area for the sample household that was shown in Figure 8.1 and Figure 8.2. The Modified Kernel Density area for this household is 77.1 sq Km.

Comparing figures 8.1, 8.2 and 8.3, it can be seen that the Modified Kernel Density captures all locations visited by the household in March 2005 either during activity participation or driving. From figure 8.1 it can be that Confidence Ellipse area includes large swathes of area below I-20 and north of the perimeter where there are no activities by the household. In figure 8.2 the Kernel Density captures all the activity locations but does not account for the dispersion of those activities and the different routes that were used to reach those activity locations. The Modified Kernel Density method captures all activities of the household, and the spatial dispersion of the activities in terms of the route instead of the linear distance between activity locations. The Modified Kernel Density method may represent the spatial appetite of the household in the activity participation model better than the Confidence Ellipse and Kernel Density methods.

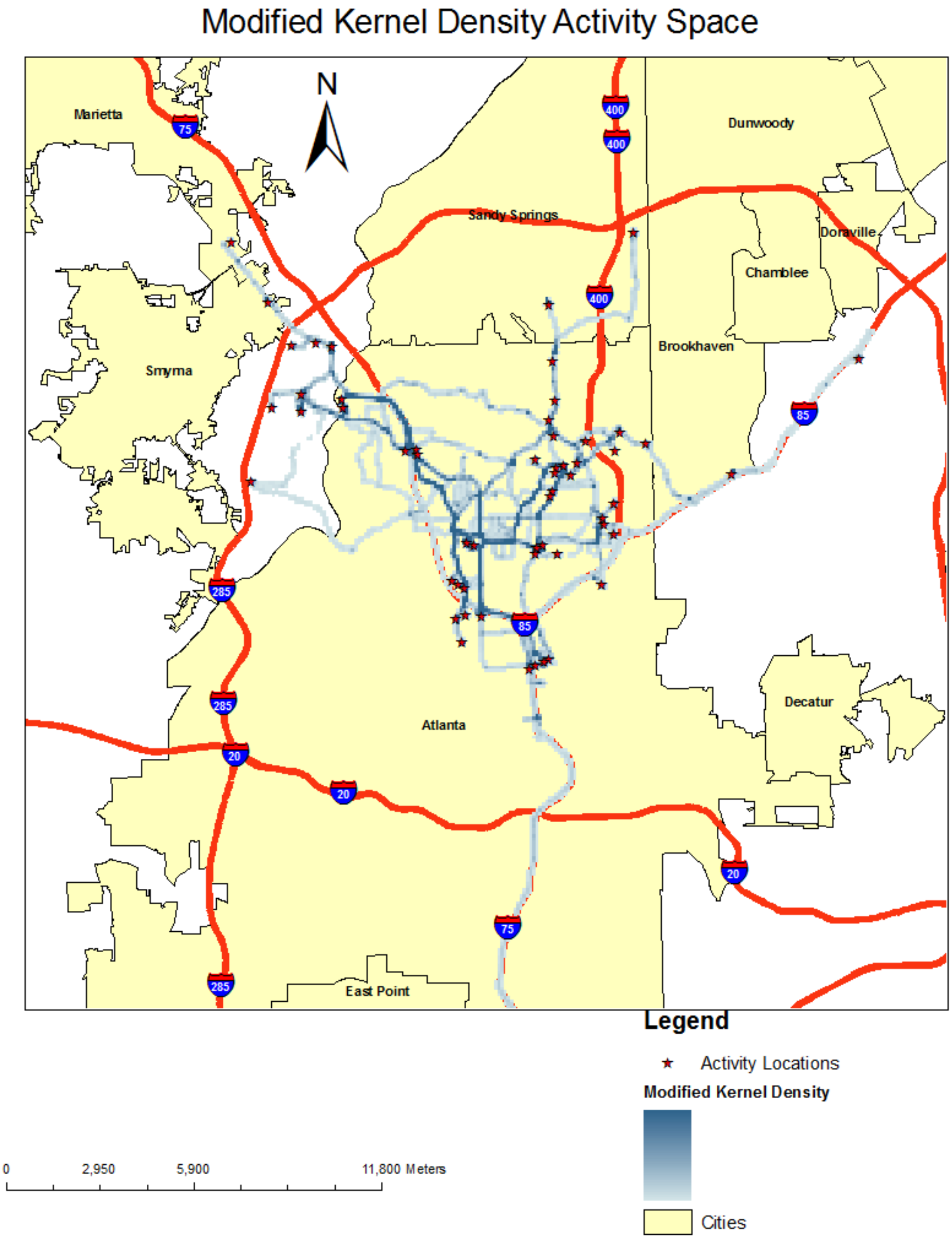


Figure 8.3 Modified Kernel Density Area Generated for Sample Household

Summary

The research effort explored the methods to estimate activity space and the use the activity spaces in travel behavior. The standard Confidence Ellipse method is good at estimating the spatial extent of the activity space (especially for general comparative purposes) but does not reflect changes in the activity frequency and incorporates space that the driver is not aware of and will never interact with. The Kernel Density estimates the effect of activity frequency on the activity space. However, the Kernel Density method has limitation in reflecting the extent of the spatial activity. A new methodology, Modified Kernel Density area is proposed in this research. The Modified Kernel Density area includes the Kernel Density area of trip end activities as well as the Kernel Density area of the travel activity. The method considers the travel between activities as an activity in itself and includes that in the measure. This measure is expected to better reflect the extent of the spatial activity as well as the frequency of the activity. An accurate measure of the activity space can better explain the spatial appetite of the household and the interactions of the spatial appetite with activity participation in the modeling process.

CHAPTER 9

ACTIVITY SPACE ESTIMATION RESULTS

This chapter presents the results for the application of the three activity space estimation methods developed in Chapter 8, namely the Confidence Ellipse, Kernel Density, and Modified Kernel Density. The first section explores the relationships between the activity space estimates and the socio-economic variables available for the Commute Atlanta data. The second section studies the correlation between the activity space estimates and other travel behavior estimates, such as number of trips and distance traveled. The next section implements a case study to assess changes in the activity space estimates with changes in household demographics. This analyses in this section seek to better understand the capabilities and limitations of each method with respect to capturing changes in travel behavior. The last section summarizes the results presented in this Chapter.

Activity Space Estimates and Socio Economic Variables

The analytical work presented here uses regression tree methods to assess the interactions between the activity space calculated by the three methods described in Chapter 8 and household demographics. Regression trees help to quickly assess potential relationships between variables within data sets. Tree-based methods highlight the variables that most significantly influence data variability [57, 82-85]. The dissertation used regression tree methods to analyze the data because the methods are easy to use and are effective for quickly identifying the demographic factors that may be related to activity space. The results of the analysis are presented in this section.

Confidence Ellipse Analysis

Figure 9.1 shows the regression tree of the Confidence Ellipse area (km-sq) by household demographics. At each node, data that satisfy the condition falls to the left and the rest of the data fall to the right. When entered into the tree analysis, the factor that explains the most variability in Confidence Ellipse is the vehicle ownership. Households that own four or more vehicles have almost twice the Confidence Ellipse area as households with less than four vehicles. Hence, the land area that encompasses 95% their monthly trips is nearly twice as large as that of households with fewer than four vehicles.

Table 9.1 summarizes the leaves of the regression tree that at least have a bin size of 18 household months since 18 months of data from the Commute Atlanta Study were used in this study. Further analysis through the tree branches and Table 9.1 indicates that households with four or more vehicles, incomes greater than $100K, and with children have large Confidence Ellipse areas as compared to other households. This result is intuitive since the above type of household has significant disposable income, large number of vehicles to participate in activities, and the presence of children in the household who require additional activities.

For households with fewer than four vehicles, household size is the next most important factor. Number of workers also plays a part in households that have two or one persons. Number of children and number of workers impact the Confidence Ellipse area of households that have more than two persons. Working, one person, low income households with vehicle ownership less than 3 vehicles have the least activity space of all demographic groups. This is intuitive since the individual in such a household have less disposable income and disposable time for non-routine activities.

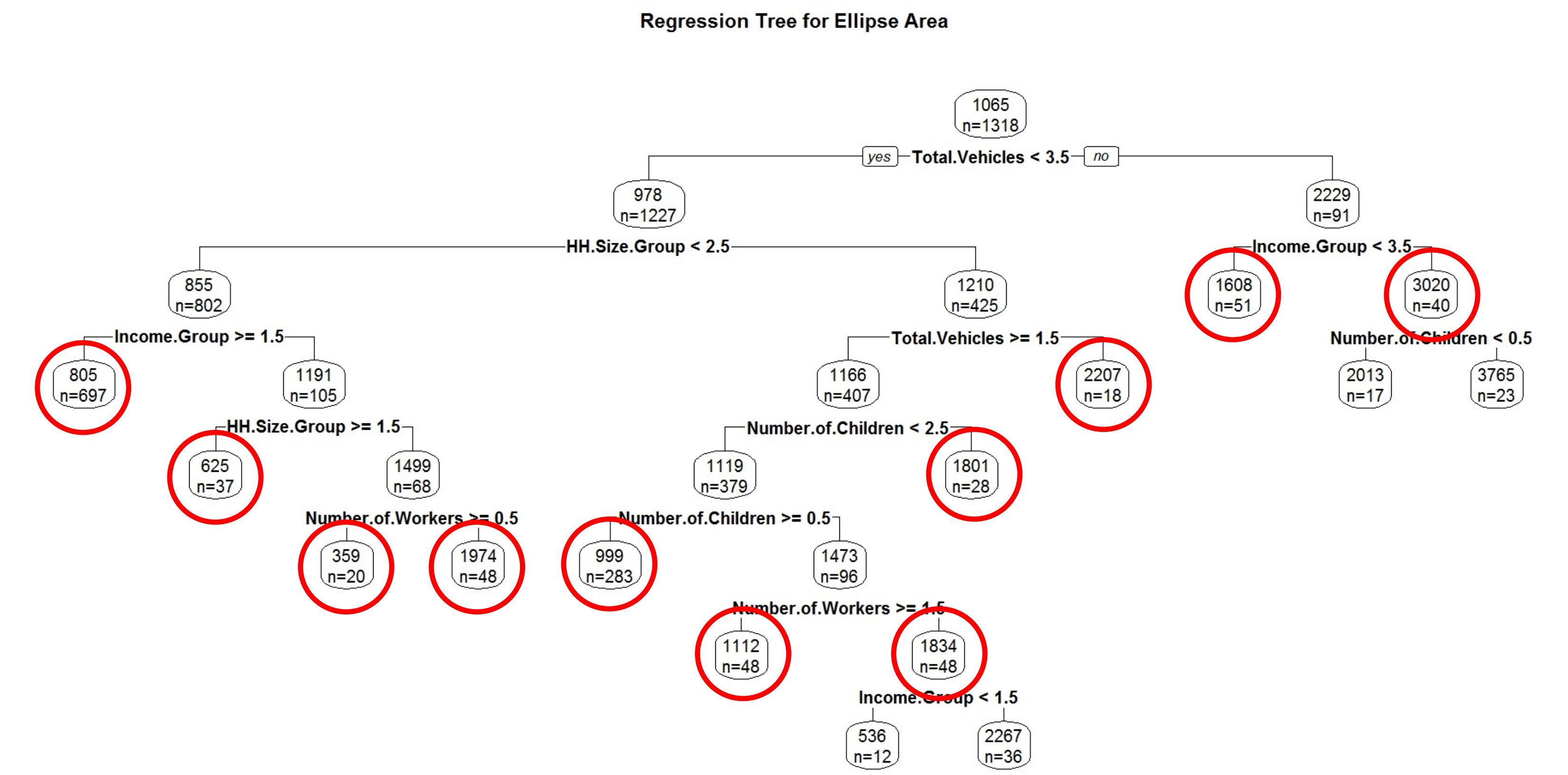


Figure 9.1 Household Demographics and Confidence Ellipse Area

Table 9.1 Summary of Regression Tree Leaves for Confidence Ellipse

<table>
<tr><td rowspan="2">Total Vehicles >3</td><td colspan="5">Very High Income Group</td><td>3020</td></tr>
<tr><td colspan="5">Other Income Groups</td><td>1608</td></tr>
<tr><td rowspan="9">Total Vehicles <=3</td><td rowspan="5">Household Size >2</td><td colspan="4">Total Vehicles =1</td><td>2207</td></tr>
<tr><td rowspan="4">Total Vehicles >1</td><td colspan="3">Number of Children >2</td><td>1801</td></tr>
<tr><td rowspan="3">Number of Children<=2</td><td colspan="2">Children Present</td><td>999</td></tr>
<tr><td rowspan="2">No Children</td><td>1 or no Workers</td><td>1834</td></tr>
<tr><td>2 or more Workers</td><td>1112</td></tr>
<tr><td rowspan="4">Household Size <=2</td><td rowspan="3">Low Income</td><td rowspan="2">One Person Household</td><td colspan="2">Non Worker</td><td>1974</td></tr>
<tr><td colspan="2">Worker</td><td>359</td></tr>
<tr><td colspan="3">Two Person Household</td><td>625</td></tr>
<tr><td colspan="4">Other Income Groups</td><td>805</td></tr>
</table>

From Table 9.1, among households with more than 2 people, the sub-group with only one vehicle have a larger activity space than households with higher vehicle ownership. This is counter-intuitive since we expect if all other demographics are controlled, higher vehicle ownership to result in larger activity space. On exploring further there is only one household that falls into the single vehicle ownership sub-group. The spatial appetite of the household may influence the Confidence Ellipse area size and probably explains this anomaly.

Kernel Density

Figure 9.2 shows the regression tree results of Kernel Density area by the household demographics. The household vehicle ownership is again the most important factor. However, the first split in this tree is for households that own three or more vehicles and households that own less than three vehicles (whereas the tree analysis using the Confidence Ellipse area split at four vehicles/household). Table 9.2 summarizes the leaves in the Kernel Density regression tree that have a minimum bin size of 18.

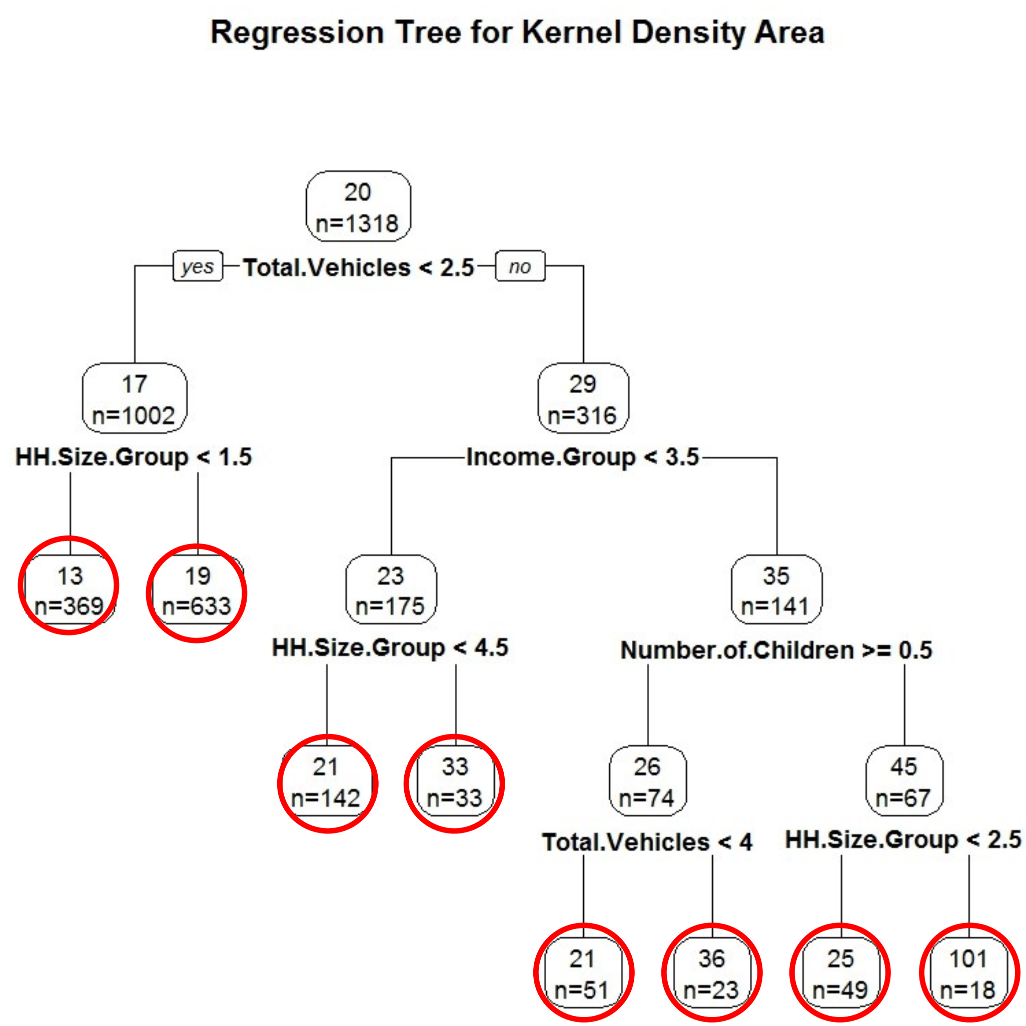


Figure 9.2 Household Demographics and Kernel Density Area

Table 9.2 Summary of Regression Tree Leaves for Kernel Density

<table>
<tr><td rowspan="2">Total Vehicles <=2</td><td colspan="3">Single Person Household</td><td>13</td></tr>
<tr><td colspan="3">Multi Person Household</td><td>19</td></tr>
<tr><td rowspan="6">Total Vehicles >2</td><td rowspan="4">Very High Income Group</td><td rowspan="2">No Children</td><td>Household Size >2</td><td>101</td></tr>
<tr><td>Household Size <=2</td><td>25</td></tr>
<tr><td rowspan="2">Children Present</td><td>Total Vehicles >=4</td><td>36</td></tr>
<tr><td>Total Vehicles <4</td><td>21</td></tr>
<tr><td rowspan="2">Other income Groups</td><td colspan="2">Household Size >4</td><td>33</td></tr>
<tr><td colspan="2">Household Size <=4</td><td>21</td></tr>
</table>

Households that own three or more vehicles, have income $100K+ and that have no children have the largest Kernel Density area. Kernel Density area correlates strongly to the number of unique locations visited by the household when the activity radius is small and the activity locations are spatially separated by large distances. While households with children make more trips and travel more miles, the number of unique locations they visit maybe more limited than their counterparts. The results of the regression tree analysis for Kernel Density area suggest that when enough resources (vehicle ownership, income etc.) are available, households without children tend to visit more unique locations than households with children.

The regression tree also indicates that one-person households with fewer than three cars have the smallest Kernel Density area which is intuitive. The findings from Table 9.2 will be compared with the results from the Confidence Ellipse regression tree results in the next section.

Modified Kernel Density Area

Figure 9.3 shows the regression tree of the Modified Kernel Density area method presented in Chapter 8 across household demographic variables. Again, total vehicle ownership is the most significant factor in explaining the variability in the Modified Kernel Density area. Similar to the Kernel Density regression tree, the split is between household with fewer than three vehicles and households with three or more vehicles. Table 9.3 summarizes the leaves of the Modified Kernel Density regression tree that have a minimum bin size of 18.

For households that have three or more vehicles, number of workers is the next factor that affects the activity space significantly. Households that have three or more vehicles, have two or more workers, and have income more than $100K have the highest Modified Kernel Density area. Among these households the sub-group of households that have only three persons has the larger Modified Kernel Density area than households with more than three people which is counter-intuitive. On further exploration there is only one household in the sub-group of three person household indicating that the spatial appetite may play a role in explaining this large Modified Kernel Density area in that sub-group. The results of this regression tree suggest that the Modified Kernel Density area mostly conforms to our intuitive expectation for demographic groups with high incomes, large vehicle ownership and large household sizes to have the largest number of trips, vehicle miles of travel and spatial extent.

Households with fewer than three vehicles and have only one person in the household have the smallest Modified Kernel Density area. In households that own more than two vehicles, the role of number of workers appears to be important. Among this demographic group, households with one or no worker has smaller spatial extent than households that have two or more workers. Among households with more number of workers, the higher number of trips, and the larger distances traveled are expected to contribute to the Modified Kernel Density area and the regression tree confirms this.

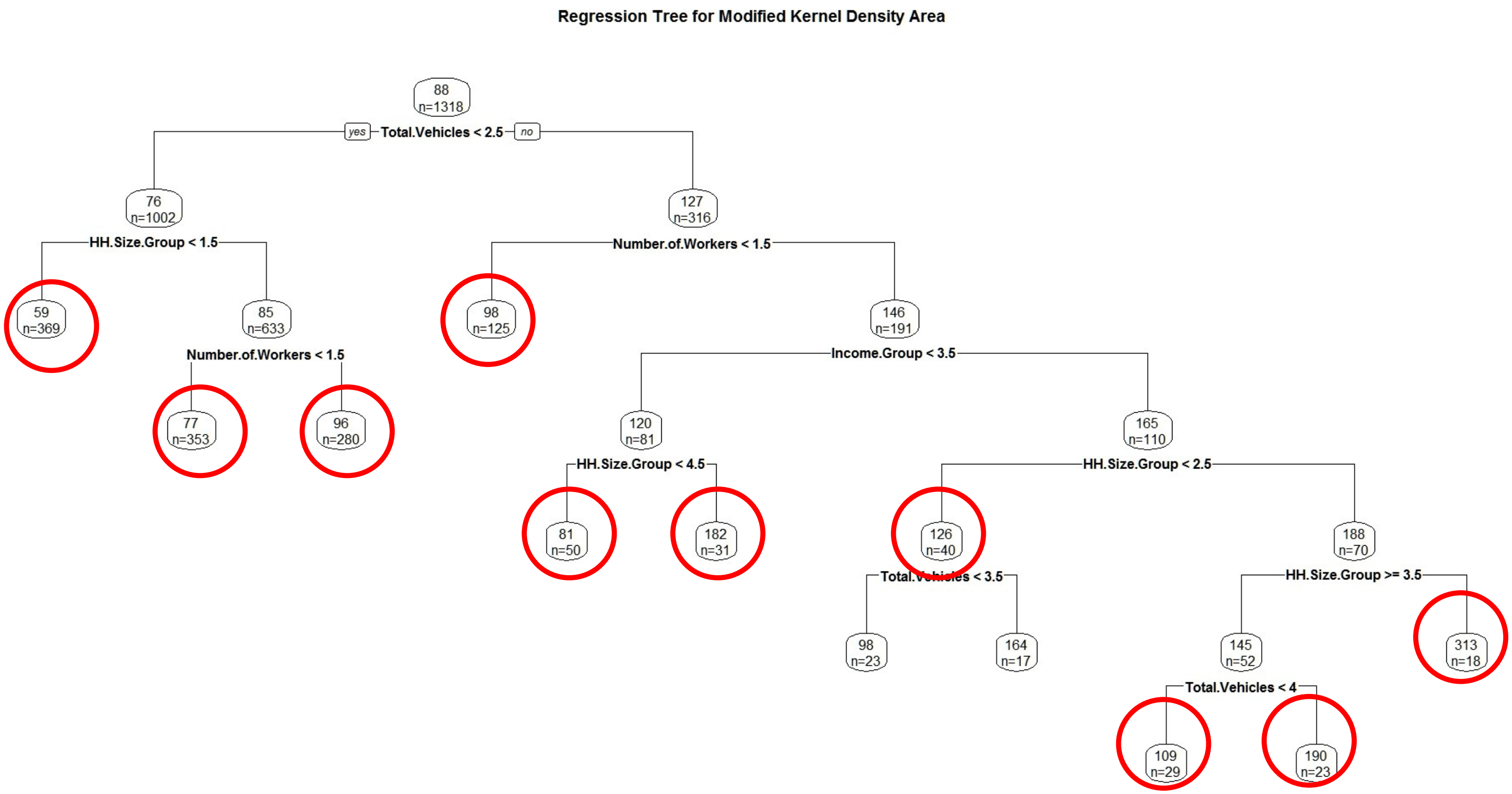


Figure 9.3 Household Demographics and Modified Kernel Density Area

Table 9.3 Summary of Regression Tree Leaves for Modified Kernel Density

<table>
<tr><td rowspan="2">Total Vehicles <=2</td><td colspan="5">Single Person Household</td><td>59</td></tr>
<tr><td rowspan="2">Multi Person Household</td><td colspan="4">One Worker or No Workers</td><td>77</td></tr>
<tr><td></td><td colspan="4">Two or more workers</td><td>96</td></tr>
<tr><td rowspan="7">Total Vehicles >2</td><td colspan="5">Single Worker or No Worker</td><td>96</td></tr>
<tr><td rowspan="6">Two or more Workers</td><td rowspan="4">Very High Income Group</td><td rowspan="3">Household Size >2</td><td colspan="2">Household Size =3</td><td>313</td></tr>
<tr><td rowspan="2">Household Size >3</td><td>Total Vehicles >=4</td><td>190</td></tr>
<tr><td>Total Vehicles <4</td><td>109</td></tr>
<tr><td colspan="3">Household Size <=2</td><td>126</td></tr>
<tr><td rowspan="2">Other Income groups</td><td colspan="3">Household Size >4</td><td>182</td></tr>
<tr><td colspan="3">Household Size <=4</td><td>81</td></tr>
</table>

Activity Space Method Results across Demographic Variables

This section compares the results of the regressions tree analysis for the Confidence Ellipse, Kernel Density and Modified Kernel Density methods. Household vehicle ownership has the highest impact on the activity space calculated by the three methods. In the Confidence Ellipse method and the Kernel Density method, household size and very high income have the next most significant effect whereas in the Modified Kernel Density method, the household size and the number of workers have the next most significant effect. This can be explained by the presence of the travel activity in the Modified Kernel Density which is expected to be affected by the number of workers. The Confidence Ellipse and the Kernel Density methods only account for the dispersion of activity locations or the number of activity locations. From the results it may be inferred that very high income impacts activity locations and their dispersion whereas the number of workers impact the amount of travel activity which is intuitive.

The number of children variable appeared in the Confidence Ellipse and the Kernel Density methods whereas it did not have a significant impact on the Modified Kernel Density method. This suggests that while number of children significantly affects the number of activity locations and their dispersion, its impact on the distance traveled and number of trips (which impacts the Modified Kernel Density area) is not more significant than the household size, high income or number of workers.

Standard recruitment procedures for household travel surveys use household income, number of vehicles, and household size to stratify their samples [17]. While these variables are important for capturing the activity space of the travel behavior, number of workers and number of children also play an important role in the activity space. While travel behavior is usually quantified in terms of number of trips and distance traveled, the activity space within which the travel occurs may provide valuable information on where the travel activities occur that may be used in recommendations for future transportation planning. The dissertation recommends that travel survey recruitments should explore the use of number of children and number of workers in their sample stratification.

Correlation between Activity Space and Number of Trips/ Trip Distance

The data exploration with the regression tree provides results for all the three methods that are consistent with expected results given the benefits and limitation of the methods. The next step is to analyze the correlation between estimated activity space and number of trips and vehicle miles of travel. As discussed in Chapter 7, trips per day (or month) and daily vehicle miles of travel by each household are good quantitative surrogates for overall travel behavior and directly impact the activity space for a household. The correlation between these variables and the activity space area will indicate the potential effectiveness of each methodology in modeling travel behavior and activity space.

Table 9.4 presents the Pearson correlation coefficient between number of trips and the activity space estimates. All of the correlation coefficients are statistically significant. The activity space area from the Confidence Ellipse methodology poorly correlates to the number of trips. This result conforms to our expectation since Confidence Ellipse captures the spatial extent of activities better than the frequencies and number of activity locations. The Kernel Density area has the highest correlation. The number of trips is directly proportional to the number of locations visited and Kernel Density effectively captures that. The correlation of the Modified Kernel Density measure is essentially equivalent to the Kernel Density measure and the additional contribution of roadway activity to the Kernel Density has not significantly improved the performance of the modified measure in reflecting the number of activity locations. In fact, the slight decrease in the correlation coefficient is likely the result of dilution of the trip end elements in the standard Kernel Density method through the addition of the roadway activity. However, it will be seen in the next section, the addition of the roadway activity to the analytical procedure significantly improves the correlation between the Modified Kernel Density method and daily miles of travel.

Table 9.4 Pearson Correlation Coefficient between Number of Trips and the Three Activity Space Estimates (N=1318)

	Confidence Ellipse	**Kernel Density**	**Modified Kernel Density**
Pearson Correlation Coefficient	0.201	0.786	0.765
Significance	0.000	0.000	0.000

Table 9.5 presents the Pearson correlation coefficient between vehicle miles of travel (VMT) and the three activity spaces. The Confidence Ellipse correlates with the VMT better than it correlates with the number of trips. However, the Confidence Ellipse is still not correlated as well with the activity parameters as either of the two Kernel Density activity space estimates. The correlation between the standard Kernel Density area and VMT is lower than observed for trips/day presented earlier. This is expected because VMT is influenced both by the number of locations and their spatial extent. However, the Modified Kernel Density area strongly correlates with the VMT and the correlation coefficient is close to the one for the number of trips. This is not surprising because the addition of the roadway activity elements to the Kernel Density calculation better reflects total travel area than the focus on destinations alone.

Table 9.5 Pearson Correlation Coefficient between Vehicle miles of travel and the Three Activity Space Estimates (N=1318)

	Confidence Ellipse	**Kernel Density**	**Modified Kernel Density**
Pearson Correlation Coefficient	0.495	0.567	0.749
Significance	0.000	0.000	0.000

Based on the correlation analyses presented above, the Modified Kernel Density appears to model activity space area that is consistent with the travel behavior observed by number of trips and vehicle miles of travel. Use of the Modified Kernel Density approach does not appear to have significant reduction in correlation with trips/day compared to the standard Kernel Density approach. Hence, the travel area identified using the Modified Kernel Density approach may be the best measure to use for activity participation modeling.

Case Study Example

The next step is to apply the three methods to a household that experienced demographic changes during the course of the Commute Atlanta study and analyze the results of the activity space with respect to other travel behavior parameters such as number of trips and vehicle miles of travel.

A household with six persons, two workers and two children that experienced vehicle ownership and income change between March 2005 and March 2006 is used for the case study. This household's income increased, moving the household from the "$75K-$100K" income group into the "more than $100K" income group and the vehicle ownership increased from three to five. The activity space estimates before and after the changes are analyzed here.

The activity space for the months of March 2005 and March 2006 were calculated by using the above three methods and compared for before and after effects of changes in demographics on travel behavior. The number of trips in March 2005 was 462 and vehicle miles of travel was 3954 miles. In March 2006, the number of trips was 684 and vehicle miles of travel was 7074. This equals to 48% increase in the number of trips and 79% increase in VMT.

Figure 9.4 shows the comparison of the Confidence Ellipse activity space for the household between the two periods. The Confidence Ellipse size in March 2006 is much larger than in March 2005. The area under the Confidence Ellipse increased from 1476 sq Km in 2005 to 3984 sq Km in 2006 which is an increase of 170%

Figure 9.5 shows the comparison of the Kernel Density area for the household. The Kernel Density is more widespread in 2006 with 211 activity locations compared to 2005 with 117 activity locations. The Kernel Density area increased from 36.8 sq Km in 2005 to 77.9 sq Km in 2006, which is an increase of 112%.

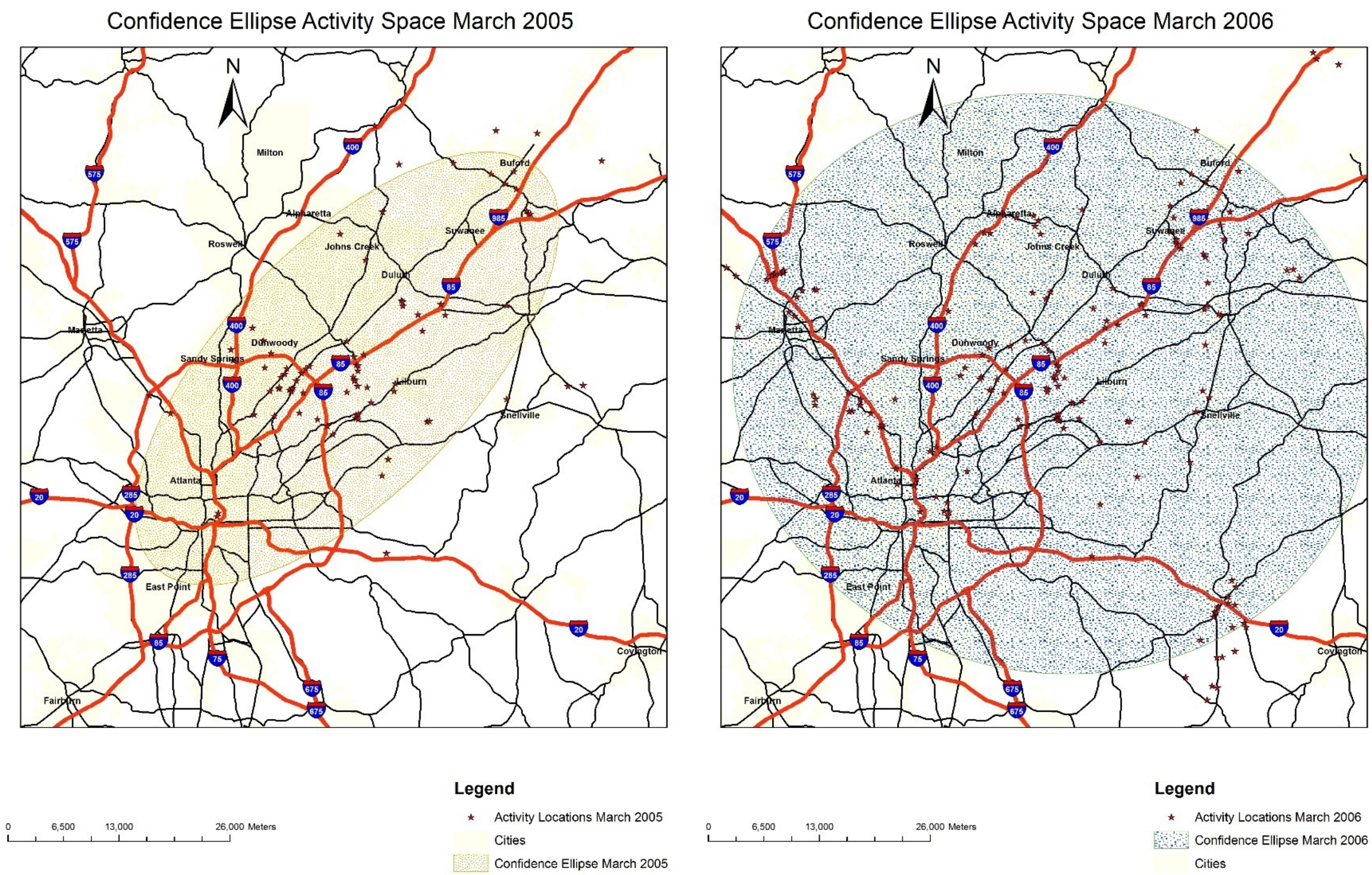


Figure 9.4 Confidence Ellipse for a Household March 2005 and March 2006

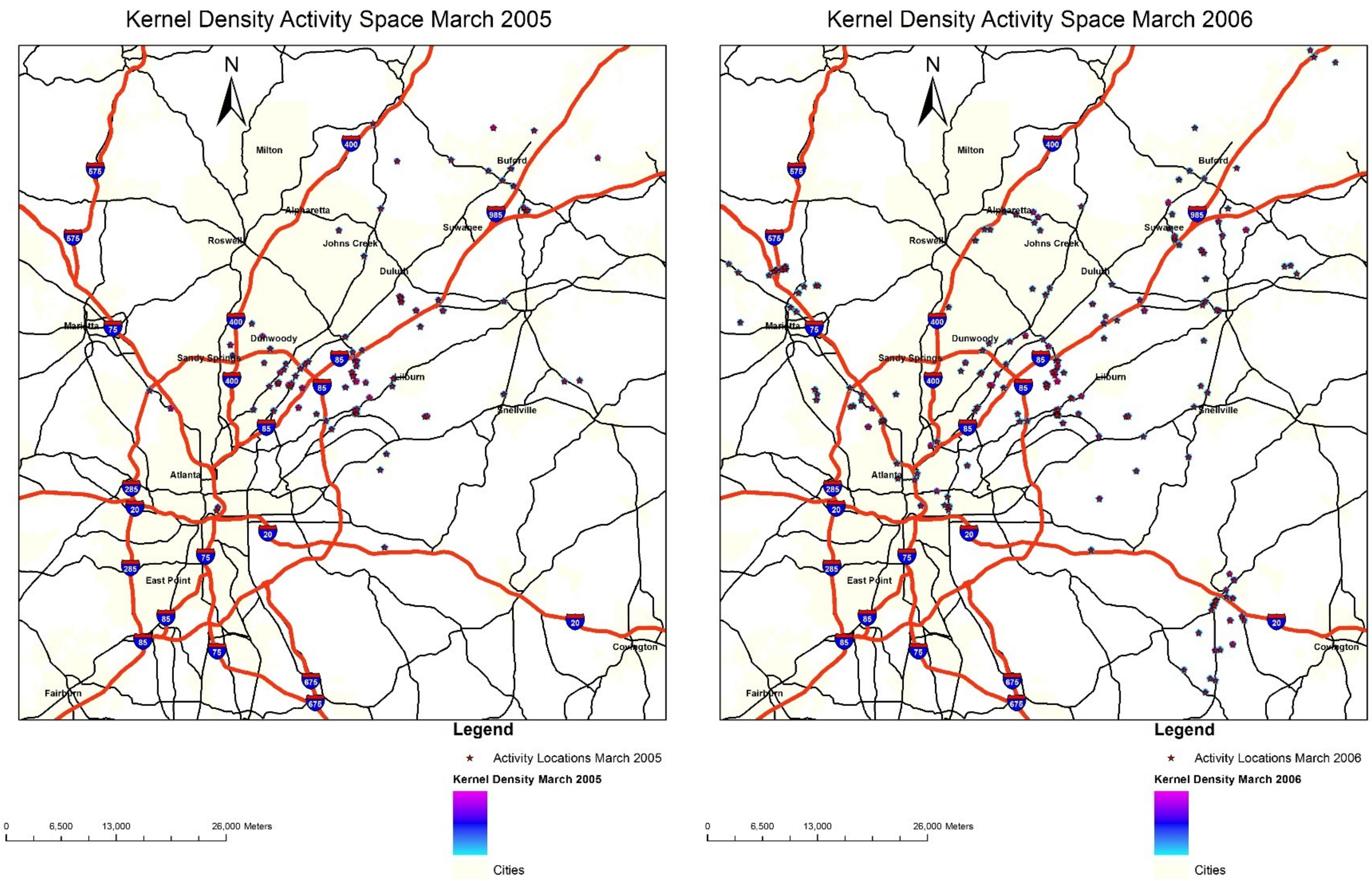


Figure 9.5 Kernel Density Area for a Household March 2005 and March 2006

Figure 9.6 shows the comparison of the Modified Kernel Density area for the household between the two periods. The Modified Kernel Density area increased from 223 sq Km in 2005 to 397 sq Km in 2006, which is an increase of 78%. Table 9.6 summarizes the travel behavior characteristics and activity space estimates of the case study household.

Table 9.6 Summary Travel Behavior Characteristics and Activity Space Estimates of Case Study Household

Measure	March 2005	March 2006	Percent Change
Total Number of Trips	462	684	48.1%
Total Vehicle Miles Traveled	3954	7074	78.9%
Number of Activity Locations	117	211	80.3%
Confidence Ellipse Area	1476	3984	169.9%
Kernel Density Area	36.8	77.9	111.7%
Modified Kernel Density Area	223	397	78.0%

In the above case study, the Confidence Ellipse and the Kernel Density methods show increase in spatial extents significantly greater than what was observed in terms of changes to number of trips, number of activity locations and vehicle miles of travel. Figures 7 through 9 indicate that the household experienced an increase in number of locations visited as well as the spatial extent of those activities between 2005 and 2006. The increase in spatial extent significantly influences the increase in the Confidence Ellipse area by 170%. However, most of the area under the ellipse in 2006 was not part of the household interaction. The increased number of locations visited increased the Kernel Density area by 112%, which was much more than the increase in number of trips, number of activity locations and vehicle miles travel. The two traditional methods appear to exaggerate the effects of changes in spatial extent and number of locations visited.

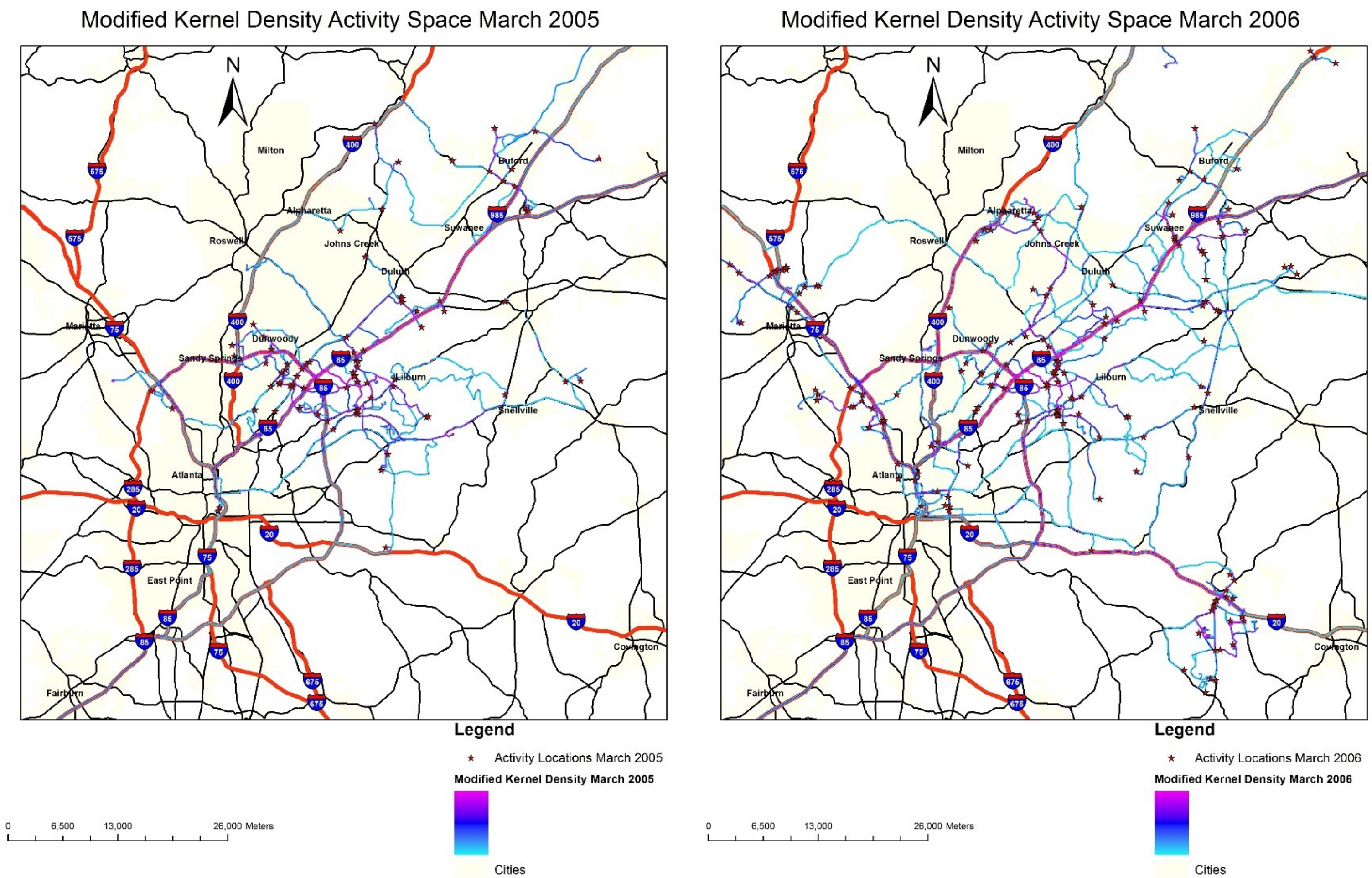


Figure 9.6 Modified Kernel Density Area for a Household March 2005 and March 2006

The change in the Modified Kernel Density from March 2005 to March 2006 is consistent with the changes in the number of activity locations and total vehicles miles traveled. The change in the Modified Kernel Density is closest to the change in the total number of trip among the three spatial activity estimates. The Modified Kernel Density method reflects the changes in travel behavior better than the Confidence Ellipse and Kernel Density methods.

Summary

The three methods were applied to the Commute Atlanta Data. Regression tree analysis was used to explore the activity spaces of the three methods with respect to household demographics. Vehicle ownership was the most significant factor affecting the activity space followed by income, household size and number of workers. The results of the Kernel Density regression tree suggest that when enough resources (vehicle ownership, income etc.) are available, households without children tend to visit more unique locations than households with children.

A correlation analysis was performed between the activity space estimated from the three activity space methods and the quantitative travel behavior surrogates; number of trips and vehicle miles of travel. The Confidence Ellipse poorly correlates with the number of trips and did marginally better with the vehicle miles of travel. The Kernel Density area correlated well with the number of trips, but slightly worse for vehicle miles travelled. The Modified Kernel Density area correlated well with both number of trips and vehicle miles of travel and was a significant improvement over the other two methods for vehicle miles of travel.

A case study was done to evaluate the effectiveness of the three methods to capture changes in activity space with changes in demographics within a household. The changes in the Modified Kernel Density area were closest to the changes in number of trips and vehicle miles of travel. Confidence Ellipse and the Kernel Density methods exaggerated the activity space estimates due to the increase in the spatial extent and the number of locations visited by the household between the two study periods. From the results of this chapter the Modified Kernel Density area has the potential to be an effective activity space measure. Therefore, the Modified Kernel Density area may be the best spatial activity measure to be used in activity participation models developed in the following chapters.

CHAPTER 10

ACTIVITY PARTICIPATION MODEL DEVELOPMENT

With the travel behavior variability and the extent of activity space being estimated the final step in the dissertation is to build activity participation models and evaluate them. Activity based travel demand modeling is based upon the premise that travel and trip making arise from the individual's need to participate in different activities that are spatially distributed [5]. A variety of activity based models in the literature have employed Discrete Choice Models, Hazard Duration Models or Structural Equations Modeling methods [5, 30, 36]. Discrete choice models are based on the allocation of time to different activities to maximize the utility for the user [30]. The multinomial logit (MNL) method has been the most common method for modeling discrete choices. Hazard duration models focus on end of duration of events given that the duration has already lasted to some specified time [5]. The conditional probability of ending an activity recognizes that the likelihood of ending an activity depends on the elapsed duration in that activity [5].

The common underlying assumption between discrete choice model and hazard duration model is utility maximization, which has been criticized for not being realistic about the way people make decisions [30]. An individual's activity participation decision making process may rely more heavily on habitual patterns than on utility maximization [30, 33, 34]. It is difficult to infer casual relationships from discrete choice model and hazard duration models because the various assumptions that are made in building the model cannot be tested [30]. Structural equation modeling (SEM) is a relatively new method used by social and behavioral scientists to test and estimate casual relationships using a combination of statistical data and qualitative casual assumptions [86, 87]. SEM can create a comprehensive framework that captures direct relationships between activity demand and the need to travel, interrelationships between the need to participate in different activities and the feedbacks from travel time to activity time [36]. SEM provides a comprehensive method for the quantification and testing of theories including casual relationships. SEMs also explicitly take into account the measurement errors that are ubiquitous and contain latent variables [86].

This chapter describes the activity participation model development. The first section briefly describes the Structural Equations modeling technique. The second section describes the potential variables and rationale for their use in studying activity participation behavior. The assumptions in modeling activity participation are then outlined. The next section describes the data preparation for the activity participation modeling. The description of the models proposed for estimation and testing is presented in the next section. The final section summarizes the chapter.

Structural Equations Modeling

Structural equation modeling can handle large numbers of variables that are specified as linear combinations of observed variables [36]. The variables may be exogenous, endogenous, or unobserved latent variables. Exogenous variables are independent variables that are not influenced by other variables. Endogenous variables are influenced by other variables and they may also influence other variables. Latent variables are unobserved variables that influence observed endogenous variables or other latent variables and indicate casual relationships [88]. A structural equation model can capture the casual influence of exogenous variables on the endogenous variables, as well as the effect of endogenous variables on each other [36]. In this modeling technique, the amount of the influence, rather than the cause and effect relationship is assumed. The amount of influence of one variable on the other is measured by the total direct and indirect effects [86]. To infer cause-effect relationships between two variables X and Y, X should precede Y, covariance and correlation should exist between X and Y, and all other causal influences should be controlled [86].

Structural equation modeling is a considered a confirmatory modeling approach, because the modeler has to construct each model in terms of the effects of the variables on each other and test the validity of the proposed model. Structural equation models are estimated by using the method of moments to minimize the difference between the model variance-covariance matrix and the population variance-covariance matrix [36, 37, 86].

The mathematical layout of the model is detailed in Chapter 2 as part of the literature review and this section will detail the methodological process of building the structural equation models. There are five steps in building a structural equation model, including: model specification, model identification, model estimation, model testing, and model modification [86, 89].

Model Specification

Model specification is the most important step in structural equation modeling process given that modeling is a confirmatory approach. Model specification should be based upon information from relevant theory and knowledge (i.e. based upon previous research). It is important to specify every parameter that needs to be estimated and the relationships between the variables that are involved in the model. The theoretical model must be well-specified if it is to be consistent with the population model. Specification errors due to inclusion and exclusion of parameters and the relationships between the variables can affect the results of the model, and thus the inferences that can be made [86]. If a hypothesized relationship between two variables is validated due to spurious correlation, the inferences from the model may not be valid.

Model Identification

Given the large number of variables associated with travel behavior and their complex interrelationships, it is important to test whether the model parameters can have a unique solution and whether they can be estimated.

Parameters in the structural equation model can either be free, fixed, or constrained. Free parameters need to be estimated, while fixed parameters have specified values. Constrained parameters are unknown, but are constrained to be equal to one or more parameters in the model. All of the parameters in the model do need to be estimated for the model to be successfully identified [86]. The number of free parameters to be estimated must be less than the number of distinct values in the variance-covariance matrix (this is called the "order condition"). Order condition is a necessary, but not sufficient, condition for model identification. If there are 'p' observed variables in a model, then the number of distinct values in the variance-covariance matrix is given by

$$p(p+1)/2 \quad [86]$$

The "rank condition" requires the algebraic determination of each parameter using the variance-covariance matrix. The rank condition is a sufficient condition for model identification, but is not easy to test in applied research [86]. There is no general and sufficient test to avoid model identification problems [36, 86]. However, there are a few steps that help improve model identification.

To avoid identification problems, either one indicator for each latent variable must have factor loading fixed to one, or the variance of the factor loading is fixed to one. The reason for imposing this constraint is to set the measurement scale for each latent variable and eliminate the indeterminacy between the variance of the latent variable and the loadings of the observed variables on that latent variable [86]. Starting with a parsimonious model with minimum number of parameters also helps in model identification [86]. Once the parsimonious model is identified other parameters can be included and tested in the subsequent models.

Model Estimation

The objective of model estimation is to estimate parameters within the specified constraints that make the variance-covariance matrix predicted by the model as similar as possible to the observed variance-covariance matrix [36]. The model estimation uses a fitting function to minimize the difference between the model variance-covariance matrix and the sample variance-covariance matrix. Some of the fitting functions used in structural equation modeling include unweighted or ordinary least squares, weighted least squares, generalized least squares, and maximum likelihood functions.

The unweighted least squares method assumes no distribution, but is scale dependent, which implies that changes in the observed variables scale will yield different solutions [86]. The generalized least squares and the maximum likelihood methods are scale-free and have desirable asymptotic properties such as minimum variance and unbiasedness [86]. The weighted least squares method needs large samples to be considered asymptotically distribution free and independent of the normality assumption [86].

The maximum likelihood estimator is the most commonly used method for model estimation in travel behavior studies [36, 37]. The maximum likelihood method assumes multivariate normal distributions. However, the maximum likelihood estimation is fairly robust against violations of multivariate normality for the sample sizes encountered in travel behavior research [36]. Multiple studies have found that a minimum sample size of 200 is needed to reduce the bias to an acceptable level while using non-normal data [90, 91]. Stevens recommended a sample size of fifteen times the number of observed variables [92], while Bentler recommended sample size to be five times the number of free parameters to be estimated [93]. Typical sample sizes encountered in travel behavior research usually satisfy the above requirements and are not biased by the multivariate normality assumption [36, 37]. In ideal situations, the asymptotically distribution-free method of estimation is preferred because of its ability to accommodate dependent variables with different distributions [37]. The analyses in this dissertation will use the maximum likelihood estimation method because the endogenous variables are continuous and the sample size is sufficiently large.

Model Testing

Model testing involves the comparison of the variance-covariance matrix implied by the model with that of the sample variance-covariance matrix [86]. Several measures have been developed to assess the goodness-of-fit of structural equations model and all are based on chi-square statistic. The model chi-square value calculated by the difference between the sample covariance matrix and the model implied covariance matrix provides the overall chi-square test statistic for rejecting a null hypothesis. The goodness-of-fit index (GFI), adjusted goodness-of-fit index (AGFI), normed fit index (NFI), and parsimony fit index (PFI) are some of the goodness fit indices that vary between 0 and 1 with values greater than 0.9 considered as good fits [86]. The root mean square error of approximation (RMSEA) is another model fit parameter whose value measures the amount of error of approximation per model degree of freedom and takes sample size into account [94]. Browne, et al., who developed the RMSEA metric state that lower value of this measure indicates a better fit [94]. In the authors subjective judgment and from their practical experience that RMSEA values below 0.05 offer the best fit and they would not employ a model whose RMSEA value is greater than 0.1 [94]. The performance of models with different number of parameters are compared using the Akaike Bayesian information criterion (AIC) while using maximum likelihood estimation [36]. The goodness-of-fit measures and the RMSEA are calculated by the following formulae [86].

$$GFI = 1 - [\chi^2_{model}/\chi^2_{null}]$$

$$AGFI = 1 - \left[\left(\frac{k}{df_{model}}\right)(1 - GFI)\right]$$

$$NFI = (\chi^2_{null} - \chi^2_{model})/\chi^2_{null}$$

$$RMSEA = \sqrt{[\chi^2_{model} - df_{model}]/[(N-1)\, df_{model}]}$$

$$AIC = \chi^2_{model} + 2q$$

Where

χ^2_{model} the chi-squared value of the model

χ^2_{null} the chi-squared value of the independence model where all covariance terms are assumed to be zero in the model (worst case)

k the number of unique values in the covariance matrix

df_{model} the degrees of freedom in the model

q the number of free parameters to be estimated in a model

With large sample sizes, it may be difficult to find a model that cannot reject the null hypothesis, simply due to the effect of the sample size. For such models, the Hoelter's critical N gives the sample size for which the chi-square value will correspond to 95 percent significance [36]. As a rule of thumb, critical N larger than 200 is considered as a good fit for such models and values below 75 are considered unacceptable fit [36]. The goodness-of-fit can help the modeler to accept or reject the proposed model, compare with competing models, and to develop alternate models based on the improvements suggested by the first order derivatives of the fitting function.

Model Modification

If the model fit is not satisfactory, the next step is to modify the model to improve the fit. The first step is to examine the parameter estimates and their statistical significance[86]. Non-significant parameter estimates can be set to zero if this is not in violation of the theory behind the model. Examining the residual matrix next provides insight into any model misspecification and potential correction [86]. Most SEM software will output modification indices that suggest the reduction in Chi-Square value by either including a parameter or defining the covariance relationship between two variables [86]. While this modification index is useful, care must be taken to ensure that the modification reflects valid theory and not spurious correlation. The modified model provides the next iteration and the above five steps are continued until a satisfactory model is developed [86].

Factors affecting Activity Participation

This section discusses the various factors that affect activity participation and how these variables are used in the building the activity participation model.

Exogenous Variables

The independent variables that are generally understood to affect activity participation are socio-demographic variables such as vehicle ownership, household income, household size, number of children, age group, gender, work status, student status, and the highest education level attained [37, 40, 41]. The day of week also influences the activities in which individuals participate [95]. Weekend and holiday activities are different from the weekday activities. Friday activities can be very different than those occurring on other weekdays. While the socio-demographic variables are independent, it is also important to acknowledge that some of these variables are correlated and even highly correlated. The household size is significantly correlated with number of children [96]. Larger households and households with higher income are likely to have higher vehicle ownership [97]. Middle aged individuals are likely to have children more than younger and older individuals [98]. Younger individuals are more likely to be students than middle aged and older individuals [99]. Individuals with college degree are more likely to be employed [100]. Figure 10.1 shows these relationships between the exogenous variables.

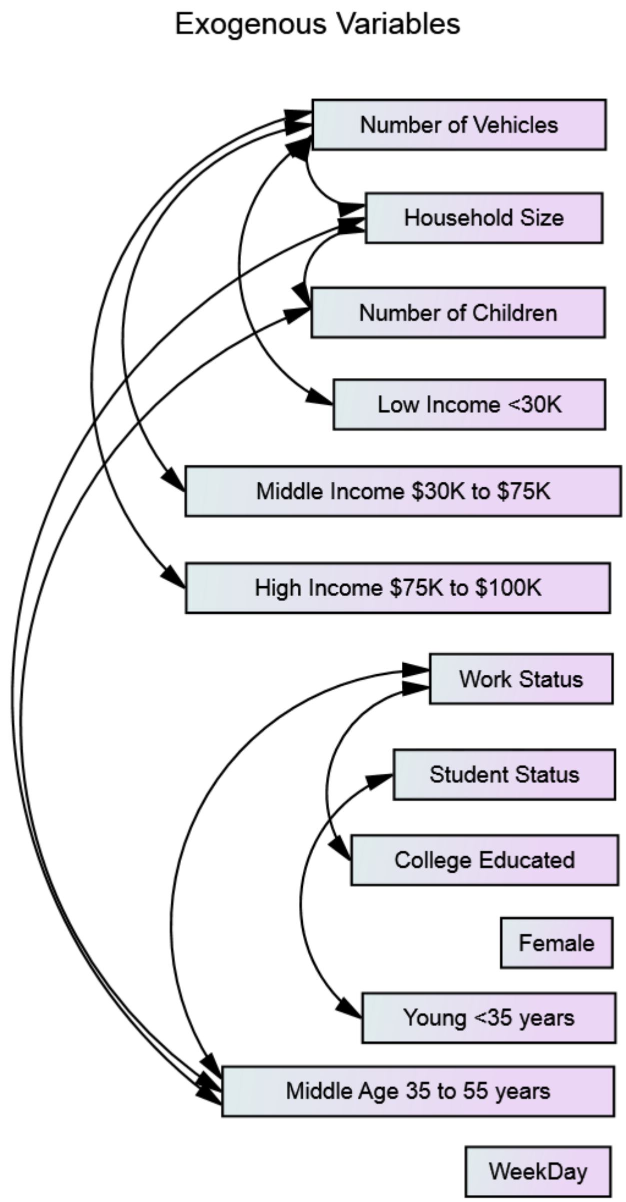


Figure 10.1 Example Covariance Relationships between Exogenous Variables

Endogenous Variables

The activity participation durations, and the travel time to these activities are dependent on the socio-demographic variables and the day of week. Activity participation durations and associated travel time durations of activities may be considered endogenous variables. All of the activity participation durations are inter-related because the total duration in a day is fixed (and travel time budgets may be fixed); hence, increasing the duration of one activity will likely decrease the participation duration of another. For the purpose of exploring the activity participation behavior potential multipurpose activity and pickup drop-off activity were not considered because both the activities were very small in number. The potential multipurpose activity also has uncertainty in the real activity that occurred and it will be hard to make inferences related to that activity.

Travel time to participate in discretionary and maintenance activities are endogenous variables influenced by socio-demographic variables. Golob et al. found a negative effect of travel duration to maintenance activities on the duration of work activities for males and a negative feedback of travel time duration for discretionary activities on the duration of discretionary activities for females [40]. The travel time to participate in home, work and school activities is and influenced by the location of those activity centers. Given the constraint that activity participation duration and travel durations must total 24 hours, the travel duration to participate in home, work and school activities were not included in the modeling process.

The activity space represented by the Modified Kernel Density, Kernel Density, and Ellipse analysis are also endogenous variables because they are likely influenced by socio-demographic variables as found in Chapter 9. The coefficient of variation of the number of trips and distance traveled are other endogenous variables. The activity space and travel variability measures are likely related to the activity participation durations of different types and are represented in the model.

Latent Variables

In behavioral sciences, we seek better understanding of relationships between two variables and try to discover if the relationship can be explained by a third variable [88]. The objective of this dissertation is to explore the possibility of using activity space measures and the travel variability measures as manifests of the spatial appetite and travel variability-seeking nature of the individual household. The "variability-seeking nature" is a latent variable that influences the travel behavior variability of number of trips per day and distance travelled. The variability-seeking nature of an individual is the appetite of an individual to participate in different activities across different days to satisfy the need for variation from routine. Schönfelder et al. found that the variety-seeking nature of the individual led to stable innovation rate or the discovery of new places even after a year in the Commute Atlanta data [3].

Assumptions

The dissertation will make the following assumptions with respect to the analysis of activity participation modeling.

- The trips made by the monitored vehicles reflect the complete travel of the household. As explained in Chapter 6, while the assumption that vehicle trips constitute the complete travel of the households misses some of the activities undertaken by the households in the Commute Atlanta study, it still captures most of the trips and activities of the households, hence, this assumption should not significantly affect the results from activity participation modeling.
- The travel time to home, work and school do not influence the activity participation in home, work and school activities (they are routine activities) and hence will not be used in the modeling. Not using the travel time to home, work and school help in accounting for the

constraint that the sum of all activity participation durations and travel time durations is 24 hours.

- The household's activity space and travel variability are used as the indicator variables for the spatial appetite and variability-seeking nature of the individual since the Commute Atlanta dataset does not include data for all participants involved in a trip (the vehicle trip is observed, but the occupants are not explicitly observed). Because the intra-household interactions are not captured in the Commute Atlanta study, it is assumed that the individual is exposed to the entire activity space and travel behavior variability of the household either by visiting those places as passengers or by interacting with household members who do. In future modeling efforts with more complete household activity data, this assumption could be specifically assessed.
- The activities at the end of each trip are associated with the primary driver of the vehicle. With no information in the Commute Atlanta study about the participants of each trip, this assumption is necessary to assign activities to individuals. The primary driver information for each vehicle was provided by the household at the time of recruitment and through follow up surveys. While this assumption may assign a trip by a different driver to the primary driver there is no way to avoid this error in the Commute Atlanta data set.
- The coefficient of variation of the number of trips and the coefficient of variation of daily distance traveled will be used as the measure of travel behavior variability as discussed in Chapter 7.
- The Confidence Ellipse area, Kernel Density area, and Modified Kernel Density area will be the activity space measures as discussed in Chapter 9.

Data Processing

The activity participation data needs to be created from the trip and daily travel summaries before it can be used for activity participation modeling. The duration of each activity type and the travel time to that activity are summed for each day and added to the daily travel summary data. Categorical variables such as income, age group, gender, work status, college status, student status, and week day are converted to dummy variables. The monthly activity space for the household and the travel behavior variability are added to the daily travel summaries. The above changes were done using Perl scripts to standardize the process, yielding a final activity participation data set.

Table 10.1 shows the description of the socio-demographic variables and their coding in the activity participation dataset. The number of household members, number of children and the vehicle ownership are interval variables and can directly be used in the modeling process. The income, gender, age group, employment status, student status, highest education level attained, and day of week are categorical variables and they are recoded by using dummy variables.

The income is represented by low income (less than $30K annual income), middle income ($30K to $75K annual income), high income ($75K to $100K) and very high income (more than $100K). The gender dummy variable is coded as 1 for ‘female’ variable and 0 for males. The employment status of a driver is presented by the ‘work status’ variable which is set to 1 if the driver of the vehicle is employed. The ‘student status’ variable represents if the driver is currently a student. The ‘College Educated” variable represents the highest level of education attained by the driver and has a value of 1 if the driver has attended some college or university and has a college degree. The age group is recoded into three variables, young (less than 35 years old), middle age (35 to 55 years old) and old (older than 55 years). The weekday variable has 1 if it is Monday through Friday and 0 if it is a weekend.

Table 10.1 Socio-Economic Characteristics Coding for Activity Participation Modeling

Variable Name	Description	Values
Household Size	Number of Household Members	1 - 6
Number of Children	Number of Children in Household	0 - 3
Vehicle Ownership	Number of Vehicles in Household	1 - 6
High Income \$75K to \$100K	High Income (dummy variable, high income =1, non-high- income =0)	0 - 1
Middle Income \$30K to \$75K	Middle Income (dummy variable, middle income =1, non-middle- income =0)	0 - 1
Low Income <30K	Low Income (dummy variable, low income =1, non-low-income =0)	0 - 1
Female	Gender (dummy variable, female =1, male =0)	0 - 1
Work Status	Employment status (dummy variable, employed =1, unemployed =0)	0 - 1
Student Status	Currently in School (dummy variable, yes =1, no =0)	0 - 1
College Educated	Highest Education Level attained (dummy variable, attended college / university =1, attended High school or less =0)	0 - 1
Young <35 years	Young (dummy variable, young =1, non-young =0)	0 - 1
Middle Age 35 to 55 years	Middle Age (dummy variable, middle age =1, non-middle age =0)	0 - 1
Weekday	Day of Week (dummy variable, week day =1, weekend =0)	0 - 1

Models for Activity Participation

This section presents the activity participation models for the drivers of the 95 households from the Commute Atlanta study. The proposed models will be estimated using structural equation models. The section proposes five models and will detail the theory behind the proposed models and how they will be compared.

Activity Participation Model - Base

The first model constructed is the base model that represents the theory of the activity based demand modeling. The socio demographic variables influence the activity participation of the individuals and the need to participate in these activities generates travel. Figure 10.2 represents the path diagram of the base model. The socio-demographic variables and the day of week are the exogenous variables that influence the activity participation durations. The endogenous activity participation durations, in turn, affect the number of trips that are generated by that individual.

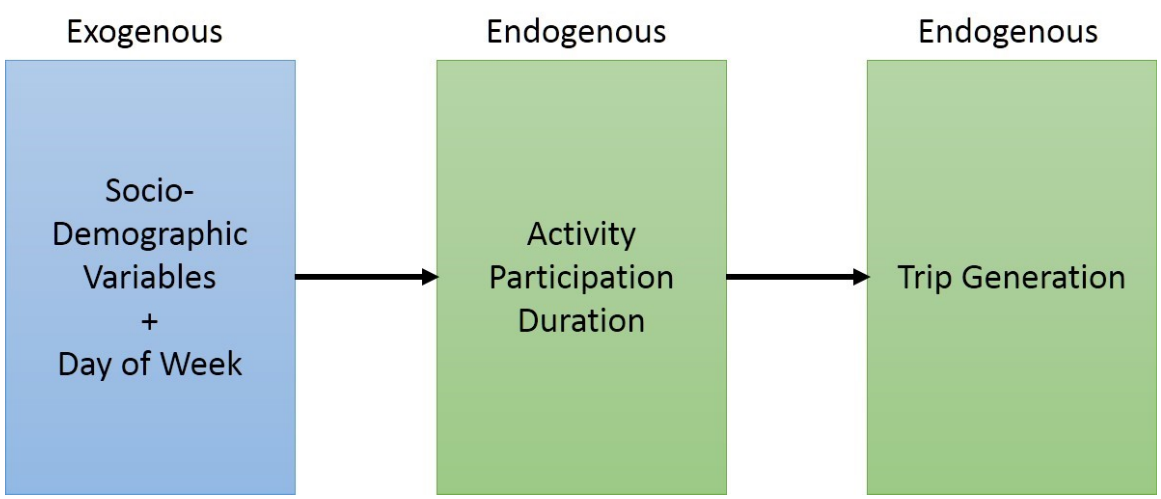


Figure 10.2 Activity Participation Base Model

Activity Participation Model with Travel Times

The next model includes travel times to maintenance and discretionary activities. Figure 10.3 represents the path diagram for this model. The travel time duration of the maintenance and discretionary activities may influence the activity participation times of those activities. If it takes longer to reach the destination for a discretionary activity, then the activity duration for that discretionary activity may be longer (a sunk-cost argument). Similarly if a maintenance activity destination needs longer travel time then it is likely that multiple maintenance activities at that destination will be combined. The travel time to maintenance and discretionary activities is influenced by the time left in a day after home, work and school activities. This model is designed to capture the effect of travel times to discretionary and maintenance activities on activity participation duration and trip generation.

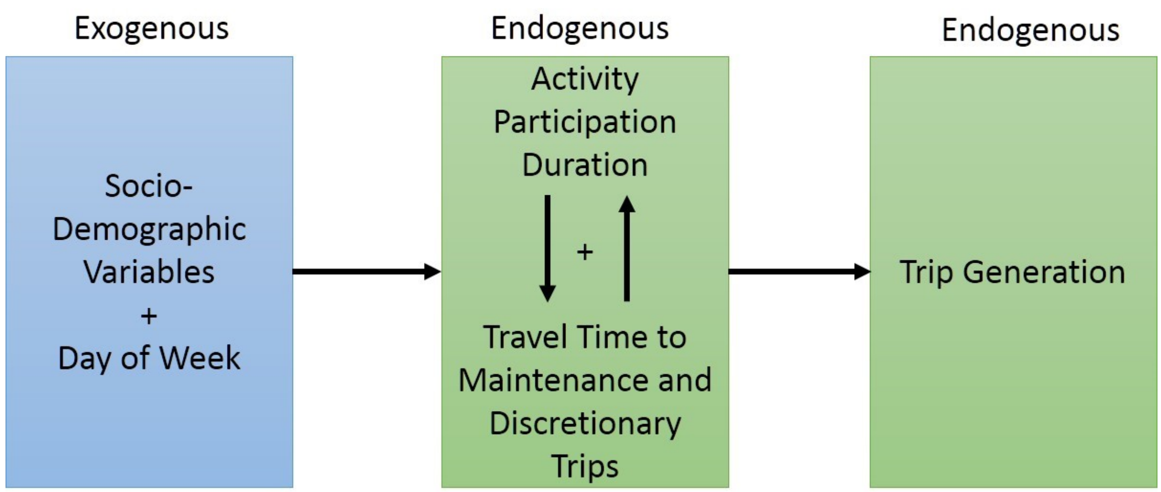


Figure 10.3 Activity Participation Model with Travel Time to Maintenance and Discretionary Activities

Activity Participation Model with Variability-seeking Nature

This third model seeks to improve on the previous model by incorporating variability-seeking nature as a latent variable which is indicated by the travel behavior variability measures coefficient of variation of trips and coefficient of variation of distance. Figure 10.4 represents the path diagram with the latent variable and the interactions between the endogenous variables. The travel behavior variables are themselves endogenous variables that are influenced by the socio-demographic variables. The travel behavior variability-seeking nature influences the individual's activity participation durations, and the travel time to maintenance and discretionary activities. The travel behavior variability indirectly affects trip generation through the activity participation durations.

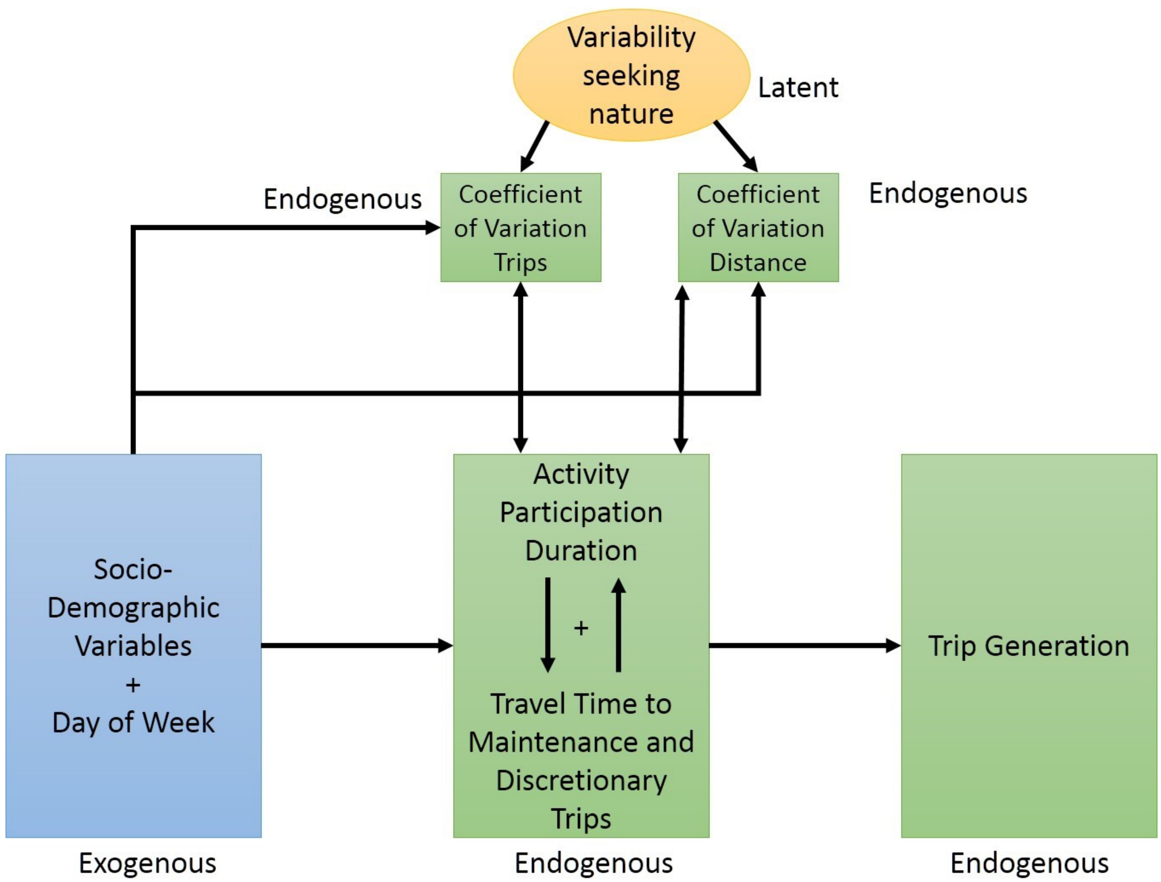


Figure 10.4 Activity Participation Model with Variability-seeking Nature Latent Variable

Activity Participation Model with Activity Space

The next model builds upon the activity participation model with travel times by adding the activity space estimates to assess their influence. Figure 10.5 represents the path diagram of this activity participation model. The activity space estimates are endogenous variables influenced by socio-demographic variables. The activity space estimates are assumed to be surrogates for the individual's appetite to travel across space. A latent variable to estimate this appetite cannot be built into the model because all the activity space indicator variables are the same measure (activity space) calculated by different methodologies. However, this model helps explore the impact of the spatial extent of the activities on trip generation.

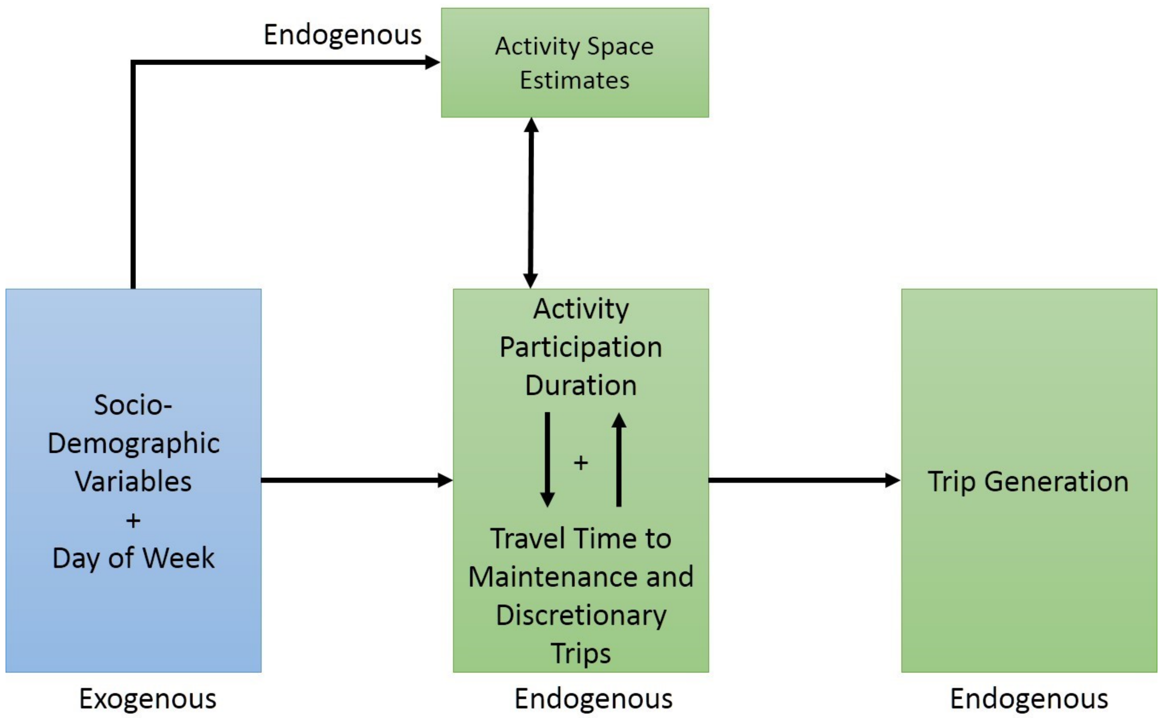


Figure 10.5 Activity Participation Model with Activity Space

Activity Participation Model with Variability-seeking Nature and Activity Space

The fifth model includes both variability-seeking nature and the activity space. The path diagram representing this activity participation model is shown in Figure 10.6. This model helps explore the combined impact of the variability-seeking nature and the spatial appetite of the individuals on the activity participation model.

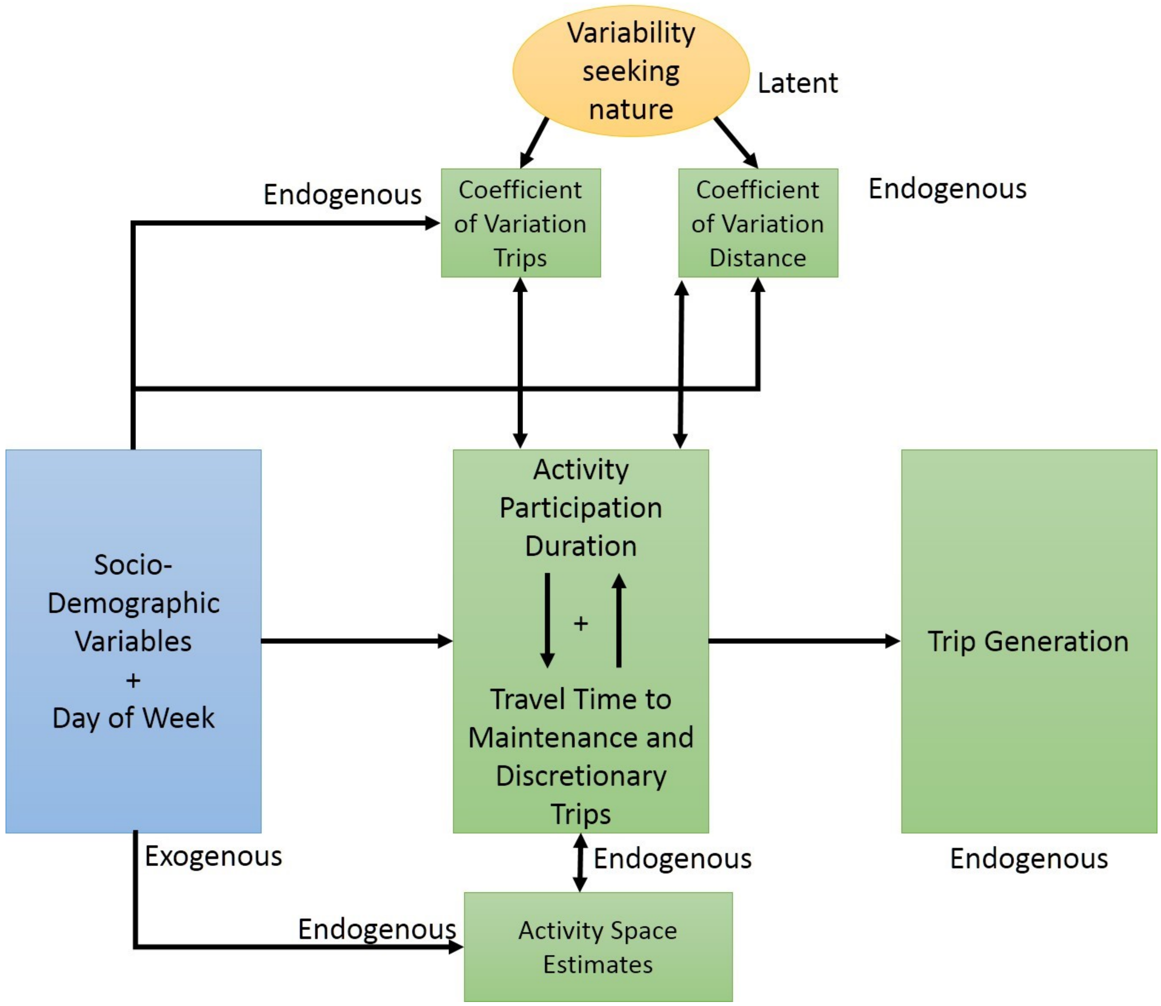


Figure 10.6 Activity Participation Model including Variability-seeking Nature and Activity Space

Model Comparison

The five activity participation models are compared in Chapter 11 using model fit parameters that can be used for different models from the same dataset. The goodness-of-fit indices, RMSEA, Hoelter's critical N and the Akaike Bayesian information criterion, all provide comparisons across models of the same dataset that have different variables. The above measures will be compared between the five models and the better model in terms of model fit and model parsimony will be identified. The role of activity space and travel behavior variability as surrogates of spatial appetite and the variability-seeking nature will be explored. The influence of the variability-seeking nature (latent variable) on the activity participation model and its role will be examined.

Summary

This chapter outlined the methodologies employed to build activity participation models. The merits of using structural equation modeling in activity participation modeling were discussed. SEM can create a comprehensive framework that captures direct relationships between activity demand and the need to travel, interrelationships between the need to participate in different activities, and the feedbacks from travel time to activity time. The steps involved in building structural equation models include model specification, model identification, model estimation, model testing and model modification. Model specification, which is the most important step, relies on prior research and theory. Model identification requires satisfaction of order condition and rank condition. The rank condition requires the algebraic estimation of each parameter in the model and the order condition requires that the number of parameters to be estimated should be less than the number of elements in the variance-covariance matrix. There are no standard tests for model identification, but certain methodological steps can increase the chances of model identification such as either one indicator for each latent variable must have factor loading fixed to one, or the variance of the factor loading is fixed to one. Model estimation includes the minimization of the difference between the model variance-covariance matrix and the sample variance-covariance matrix. Maximum likelihood fit functions are commonly used in travel behavior research to estimate structural equation models. Model testing involves evaluating various goodness-of-fit measures. Some of the measures such as goodness-of-fit indices, Hoelter's critical N and the Akaike Bayesian information criterion can be used to compare different models of the same dataset. The model modification step includes evaluating the results and modifying the specification to improve the model.

This chapter discussed the factors that influence activity participation models and identified known exogenous, endogenous and latent variables used in activity participation models. The chapter then outlined the assumptions that are made in building the activity participation models. The first assumption was that the vehicle trips constitute the entire travel for a household. The next assumption was that the travel time to home, work and school activities do not influence the durations of those activities as they are normally of a fairly fixed duration. The next assumption was that the household activity space and travel behavior are indicator variables for the spatial appetite and variability-seeking nature of the individual. The coefficient of variation of number of trips and coefficient of variation of daily distance traveled will be the travel behavior variability measures and the Confidence Ellipse area, Kernel Density Area and Modified Kernel Density area will be the activity space measures used in the modeling process.

The chapter then described the data processing of the trip and daily travel summaries to create the activity participation data. The chapter proposed five activity participation models that are estimated and compared in Chapter 11. The first model is a basic model that follows standard activity-based travel behavior modeling theory. The socio-demographic variables influence activity participation, which in turn influence the trip generation. The next model incorporates the travel times to maintenance and discretionary activities as an endogenous variable. The third model incorporates the variability-seeking nature as a latent variable, which is indicated by the coefficient of variation of the number of trips and distance traveled. The fourth model incorporates the activity space variables to explore the impact of spatial appetite of individuals on activity participation. The fifth model integrates both the variability-seeking nature and the spatial appetite into the activity participation model.

CHAPTER 11

ACTIVITY PARTICIPATION MODEL RESULTS

Chapter 11 presents the results of the activity participation modeling. The first section presents the descriptive statistics of the activity participation data. The next section provides the results of the five activity participation models proposed in Chapter 10. The following section compares the models and evaluates the use of latent variables in activity participation models. The final section summarizes the results presented in this chapter.

Activity Participation Data Descriptive Statistics

This section describes the data used in building the activity participation model. In total, 60,483 days of travel where monitored from 152 persons, driving 172 vehicles, living in 95 households. Table 11.1 shows the mean, standard deviation, and the mean of the non-zero values for each activity type. While the mean statistic provides the overall sample mean for that activity type and includes days on which the activity did not occur (therefore zero durations), it provides no insight into the actual activity duration on the days it occurred. The mean of the non-zero value activity durations provides better insight on the duration of each activity when they occurred. While the sum of all the mean activity durations is 24 hours, as expected, the sum of the mean non-zero activity durations will exceed 24 hours because different activities are undertaken on different days. The table provides the average duration of activity for occasions when the activity is undertaken (does not include zero values).

Table 11.1 Descriptive Statistics of Activity Participation Data

	N	Mean Duration (minutes)	Standard Deviation (minutes)	Non-Zero value Mean in minutes (N)
Home Activity	60483	1016.73	422.52	1090.1 (56413)
Work Activity	60483	131.13	244.69	480.3 (16512)
Maintenance Activity	60483	105.70	227.32	172.1 (37134)
Discretionary Activity	60483	96.28	269.74	284.3 (20485)
School Activity	60483	14.31	88.58	133.1 (6503)
Potential Multipurpose Activity	60483	20.39	107.27	128.64 (9590)
Pickup Drop-off Activity	60483	1.46	37.99	313.1 (282)
Travel Time to Home Activity	60483	14.71	19.57	22.1 (40172)
Travel Time to Work Activity	60483	6.18	13.94	23.3 (16075)
Travel Time to Maintenance Activity	60483	19.72	32.44	32.9 (36226)
Travel Time to Discretionary Activity	60483	8.83	24.02	28.1 (18991)
Travel Time to School Activity	60483	1.88	7.93	17.9 (6340)
Travel Time to Potential Multipurpose Activity	60483	2.61	10.54	16.9 (9310)
Travel Time to Pick-up Drop-off Activity	60483	.06	1.25	14.3 (236)
Valid N (listwise)	60483			

Summary of Activity Durations

The mean non-zero activity duration of all the activities are described in this section.

- Home activity duration was, on average for the whole sample, 18 hours and 10 minutes, which is the largest duration for any activity type. While this average home activity duration may seem high at first glance, the average incudes 11,356 days (18.7 percent of days) of no travel when individuals did not leave home. After excluding the days with no outside activity, the average home activity is about 16 hours and 10 minutes. Even this value for home activity is affected by the presence in the sample of 41 individuals who are not employed.

- Work activity - average work activity is about 8 hours for days with non-zero values, which is as expected.
- Maintenance activities, which include dining, shopping, and services, last about 2 hours and 52 minutes on average, on those days where maintenance activities occur.
- Discretionary activities last about 4 hours and 40 minutes on the days where such activities are observed. Discretionary activity includes social visits to friends and family, which helps explain the large durations for that activity.
- School activity lasts on average 2 hours and 13 minutes on those days where school trips occur. However, the Commute Atlanta Study only monitored drivers who were at least 16 years old. Therefore, school activity includes parents dropping children, adults attending universities, and vocational training, which explains the low durations for school activity.
- Pickup and drop-off activity initially averaged 6 hours and 13 minutes on those days where such activity was observed, which is much higher than expected. On further data exploration, pickup activity was noted to include airport parking for multiple days. The pickup and drop-off activity as coded was observed for fewer than half of the travel days monitored and does not seem to accurately capture the activity as intended. Therefore this activity will not be included in the activity participation modeling.

Summary of Travel Times to Activity

According to the latest Atlanta Regional Commission's travel demand model, the average home based work trip length in Atlanta is 31.6 minutes, home based other trips as 17.7 minutes, home based shopping trip length as 15.9 minutes and home based school activity as 15.5 minutes [101]. The average travel times to participate in different activities from the Commute Atlanta dataset appear reasonable for the travel characteristics in the Atlanta area. The number of days on which the travel time towards an activity occurred are slightly lower than the number of days on which the activity occurred. This is expected because the activities themselves can occasionally occur across multiple days such as visiting family for a discretionary activity. Or another example would be, if the work shift of an individual is completed at 1 AM, the trip to work would have occurred on day one and the work activity itself would have occurred over two days.

Figure 11.1 shows the distribution of the number of trips per day per person, with a mean of 4.66 trips per day over 60,483 travel days. The data set has 11,356 zero-trip days. The average number of trips per day only on the days travel occurs is 5.74 trips per day. As discussed in Chapter 6, a Tweedie distribution can be observed in the number of trips [11].

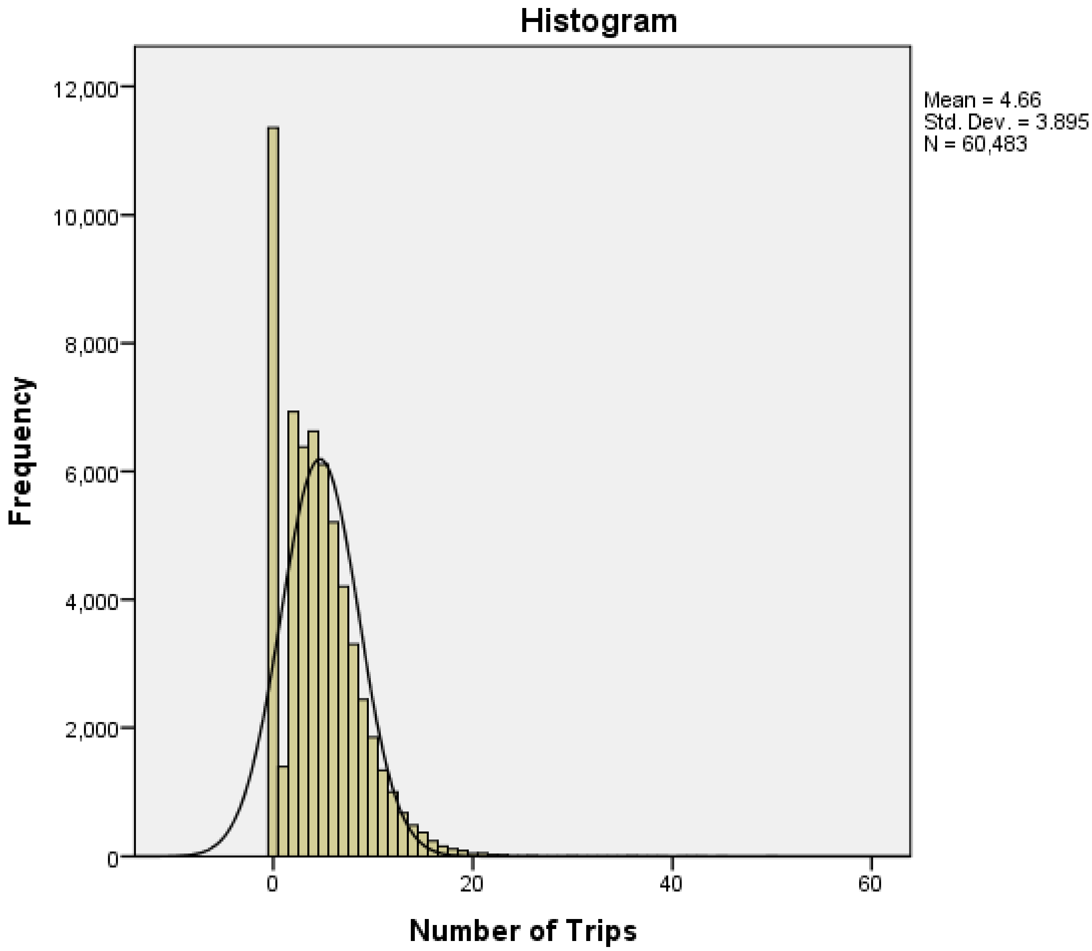


Figure 11.1 Distribution of the Number of Trips per Vehicle per Day

Activity Participation Model Results

This section presents the results of the activity participation models proposed in Chapter 10. The structural equation models representing the activity participation behavior were estimated using AMOS version 22 [102]. The AMOS software provides bootstrap resampling of the maximum likelihood estimation which is a potential solution for estimating model test statistic p values, and parameter standard errors under non-normal data conditions [103]. All the models presented in this chapter were estimated using the bootstrap maximum likelihood fit function with 1000 resampling. The model estimates, the total, direct and indirect effects, and the model fit measures are presented and discussed for each of the model.

The variables used in the modeling process as described in Chapter 10, are summarized again for the reader's convenience in Table 11.2. The driver's gender, age group, employment status, student status, highest education attained and the day of week variables are recoded into dummy variables while household size, number of children and number of vehicles owned are used as interval variables.

Table 11.2 Socio-Economic Characteristics Coding for Activity Participation Modeling

Variable Name	Description
Household Size	Number of Household Members
Number of Children	Number of Children in Household
Vehicle Ownership	Number of Vehicles in Household
High Income $75K to $100K	High Income (dummy variable, high income =1, non-high-income =0)
Middle Income $30K to $75K	Middle Income (dummy variable, middle income =1, non-middle-income =0)
Low Income <30K	Low Income (dummy variable, low income =1, non-low-income =0)
Female	Gender (dummy variable, female =1, male =0)
Work Status	Employment status (dummy variable, employed =1, unemployed =0)
Student Status	Currently in School (dummy variable, yes =1, no =0)
College Educated	Highest Education Level attained (dummy variable, attended college / university =1, attended High school or less =0)
Young <35 years	Young (dummy variable, young =1, non-young =0)
Middle Age 35 to 55 years	Middle Age (dummy variable, middle age =1, non-middle age =0)
Weekday	Day of Week (dummy variable, week day =1, weekend =0)

Activity Participation Model – Base

The base model assumes that socio-demographic variables influence activity participation, which in turn leads to trip generation, has a goodness of fit index of 0.913 and root mean square error of approximation (RMSEA) of 0.107. While the goodness of fit is acceptable, RMSEA values greater than 0.1 are not considered acceptable in the state of practice.

Table 11.3 summarizes the results of the model with the rows representing the endogenous variables and the influencing variables (both endogenous and exogenous) are in the columns. Each cell presents the total effect of the column variable on the endogenous variables. The total effect is the sum of the direct and indirect effects of the influencing variable on the endogenous variable. The direct effects indicate the regression coefficient for the influencing variable in explaining the dependent variable. The indirect effects indicate the effect of the independent variable on the dependent variable through a mediating variable which mediates the effect of the independent variable along with other independent variables on the dependent variable. For example, the indirect effect of weekday on number of trips through the mediating activity participation variables is 0.263 which means if it is a weekday the number of trips will increase by 0.263. In table 11.3, the cells that represent relationships constrained to be zero (or no relationship is hypothesized) are blank. Cells with parameter estimates that are not statistically significant are greyed. The detailed results of the activity participation models are presented in the Appendix A tables A.1 to A.15.

In table 11.3, most of the parameter estimates are statistically significant, indicating that the underlying hypothesis of interactions between socio-demographic variables and the activity participation may be reasonable. All of the parameter estimates for the impact of activity duration on the number of trips are significant and negative, which validates the hypothesis that activity participation leads to trip generation.

The home activity is significantly impacted by all of the exogenous variables, except for number of children. Household size, low income status, high income status, drivers who are students, college educated drivers, and female drivers have positive impact on the home activity duration. Vehicle ownership (number of vehicles), middle income status for the household, drivers who work, younger and middle aged drivers, and weekdays negatively impact the home activity.

Low household income does not appear to have a significant impact on the work activity duration in this model, which may be due to the presence of non-workers in the low income demographic group. Household size, number of children, high household income, drivers who are students, and female drivers have negative impact on the work activity duration. Vehicle ownership, middle income, drivers who work, younger and middle aged drivers, college educated drivers and weekdays positively impact work activity duration.

The day of week and middle income group do not exhibit a significant impact on maintenance activity duration. Household size, high income, and female drivers indicate a positive impact on the maintenance activity duration. Vehicle ownership, low income status, drivers who work, drivers who are students, college educated drivers, younger and middle aged drivers, and weekdays negatively impact maintenance activity duration.

Table 11.3 Results of base Activity Participation Model

	Weekday	Middle Age 35 to 55 years	Young <35 years	Female	College Educated	Student Status	Work Status	High Income $75K to $100K	Middle Income $30K to $75K	Low Income <30K	Number of Vehicles	Number of Children	Household Size	School Activity	Discretionary Activity	Maintenance Activity	Work Activity	Home Activity
School Activity	2.5	-1.8	-8	0	7.6	19.8	-0.3	3.5	8.5	0.9	1.9	0.4	3.3					
Discretionary Activity	-29.4	67.4	112.7	-41.5	-30.6	-25.4	-1.4	-26.6	22.9	21.9	38.2	14.5	-30.5					
Maintenance Activity	0.3	5.7	0	13.8	-18.4	-14.8	-23.6	12.2	2	-29.7	-7.3	-10.9	7.7					
Work Activity	138.9	93.1	84.1	-3.7	4.2	-29.5	110.4	-12.3	24.3	3.2	4.4	-5.3	-11.4					
Home Activity	-129	-165.7	-190	31.9	36.3	34	-101	28.5	-45.7	34.4	-36.8	0.7	32					
Number of Trips	0.263	0.112	0.09	0.01	-0.02	0.13	0.192	-0.039	-0.068	-0.33	0	-0.01	-0.01	-0.01	-0.01	-0.01	-0.01	-0.01

Work status of the driver does not appear to impact the duration of discretionary activities. Household size, high income, college educated drivers, drivers who are students, and female drivers appear to have a negative impact on the discretionary activity duration. Vehicle ownership, low income, middle income, number of children, younger and middle aged drivers are modeled as positively impacting discretionary activity.

School activity duration is not significantly impacted by work status, low income and number of children. Household size, vehicle ownership, middle income, high income, drivers who are students, college educated drivers and week day have a positive impact on school activity duration. The results suggest that school activity is not significantly impacted by number of children, which is counter-intuitive given that households with children are expected to have more school activity. The cross-tabulation of the number of children by household size on days with school activity participation data is presented in Table 11.4.

Table 11.4 Cross tabulation of Number of Children by Household Size on Days with School Activity Participation

		Number of Children				Total
		0	1	2	3	
Household Size	1	915	0	0	0	915
	2	1644	259	0	0	1903
	3	310	515	470	0	1295
	4	62	124	707	0	893
	5	10	437	478	393	1318
	6	0	0	179	0	179
Total		2941	1335	1834	393	6503

From table 11.4, a little less than half (46%) the school activities are undertaken by members of households that do not have any children. Given that the Commute Atlanta study did not monitor the activity of children, a significant portion of the school activity in the underlying data set was actually undertaken by adults attending universities or colleges. This underlying bias in the data likely leads to the non-significant impact of number of children on school activity duration.

On weekdays, the participants spend more time on school, maintenance, and work activities than they do on weekends, and spend less time in discretionary and home activities. Middle aged and younger drivers spend more time in discretionary, maintenance and work activities than in home and school activities. Females in the sample group spend more time at home, and on maintenance activities, and spend less time at work and on discretionary activities. The work status of a driver has a positive impact only on the work activity. High income drivers are spending less time in discretionary activities than are middle and low income drivers, which sound counter-intuitive. However, the lower discretionary activity duration of the high income households may also imply that high income drivers may not conduct as much social visit activity as the other income groups. The larger the number of vehicles in a household, the greater the time spent in discretionary activity, and less time is spent at home; this is expected because drivers with higher vehicle availability are more likely to make more trips.

Activity Participation Model with Travel Duration

In this second model, the travel duration to maintenance and discretionary activities are added to the previous model to enhance the activity participation modeling as discussed in Chapter 10. The goodness of fit is 0.936 and the RMSEA value is 0.098, which makes the model fit acceptable. While the RMSEA value of 0.107 in the first model is not very different from the

0.098, the comparison of the two values show the improvement in the model fit for the second model.

Table 11.5 presents the summary of the model results for the activity participation with travel time durations. The activity participation durations and the travel times to maintenance and discretionary activities exhibit a significant impact on daily trip generation. The activity participation duration have a negative impact and the travel durations have a positive impact on the number of trips. This is consistent with the hypothesis that when a driver can allocate more time to travel per day (travel duration), the driver is able to make more trips per day. For example, a one-minute increase in the travel time to discretionary activity contributes to 0.029 trips per day and a one-minute increase in the travel time to maintenance activity contributes 0.053 trips per day.

Household size, number of children in the household, and driver work status do not appear to have a significant impact on the travel time to maintenance activity. Number of vehicles, middle income, younger and middle aged drivers, and week days have a positive impact on travel time to maintenance activities. Low income, high income, college educated drivers, drivers who are students, and females have a negative impact on travel time to maintenance activities.

All of the demographic variables except student status have a significant impact on travel time to discretionary activities. Number of vehicles, number of children, work status, young and middle aged drivers exhibit a positive impact on time spent to reach discretionary activities. Household size, low income, middle income, high income, college educated drivers and females exhibit a negative impact on time spent to reach discretionary activities.

Table 11.5 Results base Activity Participation Model with Travel Time Duration

	Weekday	Middle Age 35 to 55 years	Young <35 years	Female	College Educated	Student Status	Work Status	High Income $75K to $100K	Middle Income $30K to $75K	Low Income <30K	Number of Vehicles	Number of Children	Household Size	Travel to Discretionary	Travel to Maintenance	School Activity	Discretionary Activity	Maintenance Activity	Work Activity	Home Activity
Travel to Discretionary	0	1.9	3.1	-1.6	-2.8	-0.3	0.7	-2.2	-1.5	-1.6	1.8	1.9	-1.6							
Travel to Maintenance	4.4	4.4	2.6	-2.3	-5.2	-3.7	0.5	-1.9	0.6	-4.6	0.8	-0.4	-0.1							
School Activity	2.5	-1.8	-8	0	7.6	19.8	-0.3	3.5	8.5	0.9	1.9	0.4	3.3							
Discretionary Activity	-29.4	67.4	112.7	-41.5	-30.6	-25.4	-1.4	-26.6	22.9	21.9	38.2	14.5	-31							
Maintenance Activity	0.3	5.7	0	13.8	-18.4	-14.8	-23.6	12.2	2	-29.7	-7.3	-10.9	7.7							
Work Activity	138.9	93.1	84.1	-3.7	4.2	-29.5	110.4	-12.3	24.3	3.2	4.4	-5.3	-11							
Home Activity	-129	-166	-190	31.9	36.3	34	-101	28.5	-45.7	34.4	-36.8	0.7	32							
Number of Trips	0.433	0.302	0.174	-0.13	-0.31	-0.14	0.185	-0.157	-0.031	-0.4	0.06	0.01	-0	0.03	0.05	-0	-0	-0	-0	-0

Activity Participation Model with Variability-seeking Nature

The third activity participation model enhances the previous model with the addition of the variability-seeking nature presented as a latent variable, indicated by the coefficient of variation of trips per day and distance per day. The model has a goodness of fit index of 0.938 and RMSEA value of 0.098, indicating acceptable model fits (greater than 0.9 for GFI and less than 0.1 for RMSEA). Both the model fit measures suggest that this model is an improvement over the previous two models. The model comparison will be done in detail in the next section.

Table 11.6 presents the summary estimates of the activity participation model with variability-seeking nature and their significance. The variability-seeking nature latent variable has a positive and significant impact on the indicator variables; coefficient of variation of trips and coefficient of variation of distance. The positive and statistically significant impact of the latent variable on the indicator variables support the hypothesis that the latent travel variability-seeking nature of an individual is manifested through their variability in the number of trips and the total vehicle distance traveled per day.

Table 11.6 Results for Activity Participation Model with Variability-Seeking Nature

	VariabilitySeekingNature	Weekday	Middle Age 35 to 55 years	Young <35 years	Female	College Educated	Student Status	Work Status	High Income $75K to $100K	Middle Income $30K to $75K	Low Income <30K	Number of Vehicles	Number of Children	Household Size	Coeff. Var. Distance	Coeff. Var. Trips	Travel to Discretionary	Travel to Maintenance	School Activity	Discretionary Activity	Maintenance Activity	Work Activity	Home Activity
Coeff. Var. Distance	0.19		-0.13	-0.19	0.02	0.09	0.08	-0.14	0.07	-0.05	0.08	-0.07	0.07	-0.06	0	0							
Coeff. Var. Trips	0.19		0	-0.04	0.02	-0.01	0	-0.06	0.02	0.05	0.13	-0.03	0.07	-0.07	0	0							
Travel to Discretionary	-1.2	0.02	1.89	3.06	-1.57	-2.75	-0.33	0.67	-2.23	-1.49	-1.63	1.83	1.85	-1.6	2.83	-9.29							
Travel to Maintenance	-2.91	4.39	4.32	2.39	-2.24	-5.22	-3.66	0.57	-1.88	0.65	-4.59	0.81	-0.41	-0.13	3.18	-18.9							
School Activity	-2.08	2.51	-1.75	-7.88	-0.23	7.62	19.78	-0.33	3.53	8.51	0.9	1.88	0.39	3.25	5.8	-17							
Discretionary Activity	6.55	-29.32	67.41	112.65	-41.5	-30.6	-25.4	-1.39	-26.61	22.88	21.91	38.16	14.55	-30.5	16.07	19.23							
Maintenance Activity	5.26	0.31	3.68	-4.98	14.37	-18.1	-13.2	-22.49	12.57	2.26	-28.8	-7.24	-10.7	7.91	26.15	2.21							
Work Activity	-13.9	138.74	93.39	84.88	-3.78	4.19	-29.7	110.25	-12.33	24.28	3.03	4.35	-5.37	-11.4	-52.7	-22.3							
Home Activity	13.07	-128.6	-164	-185.3	31.5	36.12	32.56	-101.7	28.22	-45.97	33.6	-36.9	0.46	31.77	-10.1	80.56							
Number of Trips	-0.23	0.43	0.3	0.17	-0.13	-0.31	-0.14	0.19	-0.16	-0.03	-0.39	0.06	0.01	-0.03	0.24	-1.5	0.03	0.1	-0	-0	-0	-0	-0

Except for student status of the driver, all other exogenous variables have a significant estimate for the coefficient of variation of trips per day. Number of vehicles, household size, college educated drivers, young and middle aged drivers have a negative estimate for the coefficient of variation of trips. All exogenous variables have a significant impact on the coefficient of variation of the distance per day. The number of vehicles, household size, middle income status, work status of the driver, and young and middle-aged drivers have negative coefficients of variation for distance per day. The gender variable does not have a significant impact on the work activity duration, whereas in the previous model (activity participation model with travel duration), gender did have a significant impact on work activity. The change in the significance suggests that the coupled variety-seeking nature of an individual and their household roles may better explain work activity duration than just their gender.

The coefficient of variation of trips per day has significant estimates for all activity participation times, except for maintenance activity. The coefficient of variation of trips per day has a positive impact on the home and discretionary activity durations, which suggests that households that seek higher variability tend to spend their activity durations either in home or discretionary activities. Drivers with high variability in the number of trips tend to spend more time at home (thereby generating lower number of trips) on some days and on other days spend more time on discretionary activities (thereby generating higher number of trips) and this alternating behavior produces the high variability. The routine activities such as work and school are not expected to be influenced by the variability-seeking nature of the individual and the model results support this hypothesis.

The coefficient of variation of the distance per day has significant estimates for all the activity participation durations. The coefficient of variation of distance per day has a negative impact on the home and work activities, and a positive impact on maintenance, discretionary, school, travel time to discretionary and maintenance activities. Again the driver's variability in the distance traveled may be due to the variations in the choice of activity locations for discretionary and maintenance activities to satisfy the travel variability-seeking nature of that individual. Home and work locations are fixed, so the higher variability in the total distance traveled indicates that variety seeking nature encourages the user to participate in more non-routine activities. The positive impact of the travel distance variability on the school activity duration maybe due to the non-routine nature of school visits such as part time students or due to the adults (who were monitored in the Commute Atlanta study) taking turns in providing rides to children in the household.

The travel variability-seeking nature has a positive impact on the coefficient of variation of trips per day, coefficient of variation of distance per day, as well as discretionary, maintenance, and home activity durations. The latent variable has a negative impact on the travel time to maintenance and discretionary activities, and duration of school and work activities. The above results imply that the desire to seek different travel routines by an individual leads to greater variation in the number of trips as well as the distance traveled per day. This desire also leads to more time spent on discretionary activities, maintenance activities, and home activities which encourage variability from the routine. The variability-seeking nature encourages the individual to spend more time in home activity on some days and spend more time on discretionary and maintenance activities on other days which produce the larger coefficient of variation of number of trips and coefficient of variation of the total distance traveled. The duration of school and work activities, which indicate more routine activity behavior, are negatively impacted by travel variability-seeking nature. The trip generation is indirectly affected by the travel variability-seeking nature and can reduce the number of trips by a factor 0.23 which implies that drivers with high variability-seeking nature on average produce fewer trips.

Activity Participation Model with Activity Space

The fourth model adds the activity space estimates Confidence Ellipse Area, Kernel Density Area and Modified Kernel Density area to enhance the second activity participation model with travel durations. The goodness of fit index is 0.941 and the RMSEA value is 0.095 both indicating acceptable model fit.

Table 11.7 summarizes the parameter estimates of the activity participation model with activity space and statistical significance of the parameter estimates. The three activity space measures, viz. Confidence Ellipse area, Kernel Density area and the Modified Kernel Density area, have significant impact on the discretionary activity, and travel time durations to discretionary and maintenance activities. The size of the activity spare area is expected to be significantly influenced by the discretionary activities and travel to discretionary activities which are non-habitual.

Table 11.7 Results from Activity Participation Model with Activity Space

	Weekday	Middle Age 35 to 55 years	Young <35 years	Female	College Educated	Student Status	Work Status	High Income $75K to $100K	Middle Income $30K to $75K	Low Income <30K	Number of Vehicles	Number of Children	Household Size	Modified Kernel Density area	Kernel Density Area	Confidence Ellipse Area	Travel to Discretionary	Travel to Maintenance	School Activity	Discretionary Activity	Maintenance Activity	Work Activity	Home Activity
Modified Kernel Density area		-7.2	0	3.5	-9.9	-1.7	17.3	-14	-28	-38.5	9.5	-25	24.2										
Kernel Density Area		-3.4	1	1.6	-2	-2.2	3.6	-6	-10	-12.4	0.6	-9.2	7.5										
Confidence Ellipse Area		450.1	281.2	-126	-290	-260	59.8	-28	19.7	152	301	-24	-2.1										
Travel to Discretionary	0	1.8	2.9	-1.5	-2.7	-0.3	0.7	-2.2	-1.5	-1.6	1.8	1.9	-1.6	0.1	-0.2	0							
Travel to Maintenance	4.4	4.3	2.3	-2.2	-5.2	-3.6	0.6	-1.9	0.7	-4.6	0.8	-0.4	-0.1	0.1	-0.1	0							
School Activity	2.5	-1.8	-8.1	-0.2	7.6	19.8	-0.3	3.5	8.5	0.9	1.9	0.4	3.3	0.1	-0.4	0							
Discretionary Activity	-29.3	67.4	112.6	-41.5	-30.6	-25.4	-1.4	-27	22.9	21.9	38.2	14.5	-31	0.2	-1.1	0							
Maintenance Activity	0.3	6.3	1.5	13.6	-18.4	-15.4	-23.9	12.1	1.9	-29.9	-7.4	-11	7.6	0.1	0.3	0							
Work Activity	138.8	93.4	85	-3.9	4.2	-29.8	110.2	-12	24.3	3	4.3	-5.4	-11	-0.6	0.9	0							
Home Activity	-129	-166	-191	32.1	36.4	34.3	-101	28.6	-46	34.6	-36.8	0.8	32	-0.1	0.3	0							
Number of Trips	0.43	0.29	0.15	-0.12	-0.31	-0.13	0.19	-0.2	-0	-0.39	0.06	0.01	-0	0	-0	0	0	0	-0	-0	-0	-0	-0

The household size does not significantly affect the Confidence Ellipse area, which represents the spatial dispersion of activity locations as discussed in Chapter 9. The household size impacts the activity frequency and the choice of locations, but not the appetite for the spatial extent of those activities and this explains the non-significant role of household size on the Confidence Ellipse area. The number of vehicles, drivers with low and middle income, work status of the driver, young and middle aged drivers have a positive impact on the Confidence Ellipse Area. All exogenous variables have significant impact on the Kernel Density area and the Modified Kernel Density area. Vehicle ownership, household size, work status of driver, female drivers and younger drivers have a positive impact on the Kernel Density area and the Modified Kernel Density area. These two variables are methodologically closer than how the Confidence Ellipse area is calculated and hence show similar trends.

The Confidence Ellipse area has no significant impact on the work and school activity, which is expected to be repetitive in nature. The Confidence Ellipse area has a positive impact on the maintenance and discretionary durations and the travel time to maintenance and discretionary activities. This is not surprising, because the Confidence Ellipse area is driven by the range of these activity locations. Discretionary and maintenance activities contribute to the number of locations visited by a driver which explains the positive associate with the Confidence Ellipse area.

The Kernel Density area and Modified Kernel Density area have no significant impact on the home and maintenance activity durations. The Kernel Density Area has a positive significant effect on only the work activity, whereas the Modified Kernel Density area has a positive effect on discretionary, school activity durations and travel time to maintenance and discretionary activities. The Modified Kernel Density appears to better explain the activity participation behavior compared to the Kernel Density, which only looks at the number of activity locations and not their extent.

Activity Participation Model with Variability-seeking Nature and Activity Space

The fifth model incorporates both the variability-seeking nature and the activity space into the activity participation model as discussed in Chapter 10. This model improves on the third and fourth models by incorporating the two measures representing the spatial appetite and variability-seeking nature into a single model. The goodness of fit index is 0.943 and the RMSEA value is 0.095 which indicate that the model is acceptable.

Table 11.8 presents the summary results of the activity participation model with variability-seeking nature and activity space. The sign and significance of the model parameter estimates are consistent with the estimates found in the previous two models, as this model combines both the model characteristics. The changes in sign of parameter estimates are associated with non-significant parameter estimates. For example the impact of Modified Kernel Density Area changed from negative in the previous model to positive in the current model. However, in both cases the parameter estimates are not significant.

Table 11.8 Results for Activity Participation Model with Variability-Seeking Nature and Activity Space

	Variability-Seeking Nature	Weekday	Middle Age 35 to 55 years	Young <35 years	Female	College Educated	Student Status	Work Status	High Income $75K to $100K	Middle Income $30K to $75K	Low Income <30K	Number of Vehicles	Number of Children	Household Size	Coeff. Var. Distance	Coeff. Var. Trips	Modified Kernel Density area	Kernel Density Area	Confidence Ellipse Area	Travel to Discretionary	Travel to Maintenance	School Activity	Discretionary Activity	Maintenance Activity	Work Activity	Home Activity
Coeff. Var. Distance	0.19		-0.13	-0.19	0.02	0.09	0.08	-0.14	0.07	-0.05	0.08	-0.07	0.07	-0.06												
Coeff. Var. Trips	0.19		0	-0.04	0.02	-0.01	0	-0.06	0.02	0.05	0.13	-0.03	0.07	-0.07												
Modified Kernel Density area			-7.22	0	3.54	-9.86	-1.7	17.31	-13.8	-27.6	-38.46	9.55	-25.1	24.16												
Kernel Density Area			-3.4	1.01	1.62	-2.03	-2.21	3.55	-5.97	-10.1	-12.42	0.62	-9.21	7.5												
Confidence Ellipse Area			450.1	281.2	-126.4	-290.4	-260.1	59.8	-27.6	19.72	152.1	301.4	-23.9	-2.11												
Travel to Discretionary	-0.7	0.03	1.8	2.85	-1.53	-2.74	-0.26	0.72	-2.21	-1.48	-1.59	1.84	1.86	-1.59	2.37	-6.14	0.09	-0.14	0							
Travel to Maintenance	-2.31	4.41	4.22	2.14	-2.2	-5.21	-3.58	0.62	-1.86	0.66	-4.54	0.81	-0.4	-0.12	3.29	-15.7	0.09	-0.08	0							
School Activity	-1.99	2.52	-1.75	-7.93	-0.46	7.63	19.77	-0.32	3.52	8.54	0.92	1.87	0.41	3.24	3.95	-14.7	0.11	-0.36	0							
Discretionary Activity	4.9	-29.25	67.39	112.6	-41.47	-30.64	-25.39	-1.38	-26.6	22.87	21.92	38.17	14.55	-30.5	13.88	12.52	0.14	-0.89	0.01							
Maintenance Activity	6.93	0.35	4.24	-3.59	14.24	-18.2	-13.64	-22.8	12.48	2.19	-29.01	-7.26	-10.7	7.84	32.85	4.53	-0.06	0.91	0.01							
Work Activity	-17.9	138.68	93.75	85.76	-3.87	4.15	-30.02	110.1	-12.4	24.24	2.86	4.33	-5.42	-11.4	-54.5	-42.2	-0.39	-0.06	0							
Home Activity	14.39	-128.8	-164.4	-186.3	31.73	36.17	32.9	-101.5	28.29	-45.9	33.78	-36.85	0.51	31.82	-15.7	93.34	0.04	-0.04	-0							
Number of Trips	-0.18	0.434	0.289	0.143	-0.122	-0.309	-0.128	0.192	-0.16	-0.03	-0.39	0.062	0.015	-0.03	0.242	-1.23	0.007	-0.01	0	0.03	0.05	-0	-0	-0	-0	-0

The magnitude of the effect of travel variability-seeking nature variable on the number of trips per day has marginally decreased from 0.23 in the activity participation model with variability-seeking nature to 0.18 in this model. Similarly the indirect effect of the activity space measures on trip generation have marginally reduced between the fourth model that had activity space measures and this model which incorporates both activity space measure and the variability-seeking nature. This is as expected because both the variables are trying to explain the same part of the trip generation and have reduced the impact of the other variable.

Discussion of Activity Participation Models

The five activity participation models presented in the preceding section represent the theory of activity based travel demand modeling by estimating trip generation as a result of activity participation. The activity participation itself is dependent on the socio-demographic characteristics, day of week, travel variability-seeking nature and spatial appetite of the individuals. These complex interactions can be effectively modeled using the structural equations modeling system. Traditional modeling techniques cannot explain the variability in travel behavior as a measure of individual's desire to experience such variability. In structural equations modeling, using observed variables that represent travel behavior variability, the influence of the individual's desire to have travel variability can be measured as was shown in the third and fifth model. The appetite for activities that are spatially distributed can be manifested using observed activity space. However, to measure the influence of the unobserved spatial appetite (latent variable), at least two different measures that indicate spatial appetite are necessary. The Confidence Ellipse area, Kernel Density area and the Modified Kernel Density area all represent the activity space, calculated using different methods and cannot capture the influence of the latent variable without another measure that indicates the spatial appetite.

Schonfelder et al. found that the individual's variety seeking nature encouraged them to visit new locations and drop other locations in a stable pattern over a year [3]. The results of the activity participation model supports that variety seeking nature in terms of spatial appetite and travel behavior variability do play a role in the activity participation of the drivers in the household. However, further research using larger datasets with better measures of spatial appetite are required to reaffirm these results. Fortunately, newer travel data collection systems using smartphones will be able to collect data across all modes and all participants in a household and help in furthering the research on the variety seeking nature of the individual.

Table 11.9 shows the comparison of the model fit parameters across the five models. The cells that have the best two model fit parameters across the five models are highlighted in green. The base model may not be acceptable under conventional practice because its RMSEA value is greater than 0.1. Among the remaining four models the final model that incorporates both variability-seeking nature and activity space has the best goodness of fit index, normed fit index and RMSEA indicating that it has the best model fit among the five models that are compared. However, picking the best model means picking the most parsimonious model that best explains the data. The Akaike information criterion (AIC) and the Hoelter's critical N both penalize models that are over-fitted. The fourth model which includes the activity space estimates into the activity participation model has the highest values of Hoelter's critical N and the second lowest AIC value. The fourth model also has the second highest goodness of fit, adjusted goodness of fit and normed fit index and the lowest RMSEA value. The fourth model of activity participation which incorporates activity space is the best model in terms of model parsimony and explanatory powers. This implies that the spatial appetite of households may be better than the variability-seeking nature in explaining activity participation and therefore trip generation.

Table 11.9 Comparison of Model fit Parameters between Activity Participation Models

Measure	Base	With Travel Duration	With Variability-seeking Nature	With Activity Space	With both Variability-seeking Nature and Activity Space
GFI	0.913	0.936	0.938	0.941	0.943
AGFI	0.797	0.815	0.795	0.798	0.782
NFI	0.853	0.892	0.896	0.92	0.923
RMSEA	0.107	0.098	0.098	0.095	0.095
AIC	56272.8	46423.44	4955.31	48291.15	50612.3
Hoelter	112	134	131	139	138

The activity participation models developed by this dissertation have considerable room for improvement. While the RMSEA is acceptable, the value of RMSEA does not indicate a great fit (requires RMSEA to be less than 0.05). The model fit may be improved by using data that more accurately captures activity participation and household interactions in the activity participation. The Commute Atlanta data used in this dissertation lacked detailed information on trip participation and household interactions. Future data collection efforts should ensure complete capture of travel by all participants across all modes within participating households. The use of additional spatial measures that indicate spatial appetite such as number of unique locations visited and number of new locations visited may better explain the activity participation models better. It is important to ensure that the resulting model parsimoniously explains the activity participation and does not just improve by over fitting.

Recommended Activity Participation Model with Modified Kernel Density Area

Based on the discussion in the previous sections, the final recommended activity participation model presented in this section incorporates only one measure of activity space and eliminates non-significant relationships between socio-demographic variables and the activity participation durations. The Confidence Ellipse area and the Kernel density area were not included in this model because the Modified Kernel Density area best explained activity participation behavior.

Table 11.10 compares the model fit parameters between the final recommended activity participation model with Modified Kernel Density area and the activity participation model with activity space. The goodness of fit index for the recommended model is 0.938 and the RMSEA value is 0.089. The AGFI, RMSEA, AIC and Hoelter's critical N are better for the model with Modified Kernel Density area than the model with all the activity spaces.

Table 11.10 Comparison of Model Fit between Activity

Measure	with Activity Spaces	with Modified Kernel Density Area
GFI	0.941	0.938
AGFI	0.798	0.838
NFI	0.92	0.897
RMSEA	0.095	0.089
AIC	48291.15	47110.8
Hoelter	139	171

The summary results of the activity participation model with Modified Kernel Density are shown in Table 11.11. The detailed parameter estimates for this model are shown in Appendix A, tables A.16 to A.18. The Modified Kernel Density Area impacts significantly all activity durations except the discretionary activity duration. All the socio-demographic variables have significant effect on the Modified Kernel Density area similar to what was observed in the activity participation model with activity spaces. The results of this model suggest that work status does not have significant impact on travel time to discretionary activities. Higher income and vehicle ownership do not appear to significantly impact travel time to maintenance activities. The household size does not appear to significantly impact the work duration of an individual.

Table 11.11 Results for Activity Participation Model with Modified Kernel Density Activity Space

	WeekDay	Middle Age 35 to 55 years	Young <35 years	Female	College Educated	Student Status	Work Status	High Income $75K to $100K	Middle Income $30K to $75K	Low Income <30K	Number of Vehicles	Number of Children	Household Size	Modified Kernel Density area	Travel Discretionary	Travel Maintenance	School	Discretionary	Maintenance	Work	Home
Modified Kernel Density area		-7.2	0.0	3.5	-9.9	-1.7	17.3	-13.8	-27.6	-38.5	9.5	-25.1	24.2								
Travel Discretionary		1.8	2.9	-1.6	-2.7	-0.1	0.7	-2.4	-1.8	-2.5	1.8	1.8	-1.6	0.1							
Travel Maintenance	4.4	4.3	2.3	-2.2	-5.2	-3.6	0.6	-2.0	0.5	-5.0	0.7	-0.5	-0.1	0.1							
School	2.5	-2.1	-8.3	0.1	7.2	19.6	0.6	2.6	7.3	-1.3	0.3	-0.8	4.6	0.0							
Discretionary	-29.5	67.2	112.7	-41.8	-30.5	-24.9		-27.2	21.6	18.3	37.5	12.8	-29.5	0.0							
Maintenance	0.0	5.9	0.0	13.5	-18.1	-14.7	-23.6	12.1	1.5	-31.2	-7.7	-12.0	8.3	0.2							
Work	138.9	93.3	83.5	-1.2	3.3	-30.6	111.2	-10.3	26.5	13.0	4.9	-5.7	-10.9	-0.3							
Home	-128.4	-165.2	-187.7	29.6	37.2	34.1	-104.0	28.5	-44.4	33.3	-34.6	5.6	28.2	-0.2							
Number of Trips	0.4	0.3	0.2	-0.1	-0.3	-0.1	0.2	-0.2	0.0	-0.4	0.1	0.0	0.0	0.0	0.0	0.1	-0.0	-0.0	-0.0	-0.0	-0.0

The female gender variable did not have a significant direct effect on the work duration in the activity participation model with all activity spaces given that the average duration spent at work by men and women was not significantly different. In this model the female gender variable has an indirect negative effect on the work duration through the Modified Kernel Density area mediating variable. This suggests that for a given Modified Kernel Density area, gender may play a role in explaining work duration; i.e. among individuals who have identical Modified Kernel Density area, men may spend more time at work than women. This model appears to explain the role of gender better than the previous models.

Summary

This chapter presents the results of the activity participation model. The descriptive characteristics of the activity participation data were explored. The travel data included 60,483 travel days from 95 households and 152 persons. The average time spent in home was about 16 hours which includes data from non-workers and weekends. The time spent at work was 8 hours on days when there was work activity. The maintenance activity was on average 2 hours and 50 minutes and discretionary activity was about four hours on the days the activities occurred. The average number of trips per day from the data set was 4.66.

The five activity participation models proposed in Chapter 10 were estimated using the AMOS software. The model estimation used bootstrapping of the maximum likelihood estimation to minimize the biases due to violation of the multivariate normality assumption. The first base model had acceptable goodness of fit but unacceptable RMSEA value. The second model incorporated the travel duration to maintenance and discretionary activities and the model had acceptable fit. The number of trips were positively influenced by the travel durations and negatively influenced by the activity durations as expected. The third model added the variability-seeking nature to the activity participation models. The variability-seeking nature latent variable had a positive effect on the travel behavior variability and it had a positive indirect effect on the number of trips. The fourth model incorporated the activity space measures instead of the variability-seeking nature. The estimates suggested that Modified Kernel Density area discussed in Chapter 9 is better than Kernel Density area in explaining the travel behavior. The fifth model incorporated both the variability-seeking nature and the activity space. This model's results were consistent with the previous two models and had a better fit index than both the models.

The five activity participation models reflect activity based travel demand modeling theory by modeling the interactions of the socio-demographic variables (exogenous) on the activity participation variables (endogenous) which impact the trip generation (endogenous). Structural equation modeling technique allows the exploration of large number of variables and complex relationships. Structural equation modeling also provide a framework to quantify the influence of the latent variables such as the variability-seeking nature on the trip generation. The comparison between the five models based on model fit should also account for model parsimony. Based on the goodness of fit measures, RMSEA, Akaike information criterion and Hoelter's critical N value, the fourth model which incorporates activity space and not the variability-seeking nature (which may represent an over-specification given the integration of activity space) was found to be the best model. This model was further refined by eliminating non-significant relationships, Confidence Ellipse area and the Kernel Density area. The final recommended activity participation mode with the Modified Kernel Density showed improvements in terms of model fit parameters and parameter estimates.

The activity participation models developed in this chapter have considerable room for improvement in terms of model fit and explanatory variables. Future data collection efforts should ensure capture of travel across all modes and across all household members to ensure a more comprehensive activity participation data to model the interactions between household members. The activity participation model can also be further enhanced by integrating other measures of spatial appetite such as number of unique locations or new locations visited by a household to measure the impact of the unobserved spatial appetite latent variable. The spatial appetite of the household plays an important role in decision making related to activity participation. Including spatial appetite more accurately in the modeling process will help in better understanding of travel behavior.

CHAPTER 12

CONCLUSIONS

The final chapter presents the conclusions and summary of this dissertation work associated with developing a methodology to model activity participation using traditional demographic parameters coupled with longitudinal travel variability and spatial activity extent. The first section presents the summary of the research findings. The second section discusses the potential contributions to research by this dissertation work. The final section examines the potential uses of enhanced spatial data and modeling tools in future planning and recommends future research efforts likely to improve activity participation modeling and understanding of travel behavior.

Research Findings

This section summarizes the dissertation work and discusses the key research findings. The main objective of this dissertation was to develop a methodology to utilize travel behavior variability and activity space as surrogates for the variability-seeking nature and spatial appetite of households in activity participation modeling. The dissertation reviewed existing literature on travel survey data collection systems, trip purpose identification methods, travel variability studies, activity space estimation methods, and activity participation modeling. The overall objective was to develop surrogates, apply the methods to enhanced spatial data, and then undertake advanced statistical methods to assess the impact of variability-seeking surrogates on travel behavior.

Second-by-second GPS travel data collected from the Commute Atlanta study (more than 1.8 million vehicle trips) were used in this dissertation effort. The raw data were processed for quality control and quality assurance, joined with household demographic characteristics, and processed to route. Trip chains were identified by examining off-road network data with speeds less than 5 mph and having a dwell time of 30 seconds or more. The data were then filtered to remove engine starts (non-trips), households with more than 10% bad GPS data, and households with inconsistent or incomplete demographic data. The research dataset includes the baseline and pricing data for 95 households for which complete travel data and socio demographic data were available (282,000 vehicle trips). Data collected by University of Minnesota during a travel behavior study on the use of the I-35W Bridge, were used to evaluate trip purpose identification methodology.

The dissertation developed a new methodology to automatically identify the purpose of a trip for passively collected GPS travel data from instrumented vehicle studies. The methodology requires the use of commercial mapping software, in this case Microsoft MapPoint, instead of using typical land use data as are often used in other studies. The methodology utilizes the habitual behavior of the users to identify regularly visited locations such as home, work, coffee shops, etc., and assign radii of parking for those locations. After assigning the trip purpose to trips that end at habitual locations, the methodology examines the business locations within a quarter mile of the remaining trips. The algorithm picks the closest businesses to the end of the trip and, depending on the types of selected businesses, assigns a trip purpose classification of maintenance, discretionary, school-daycare, pickup or drop-off, or potential multipurpose trip. The school-daycare trips are checked for time of day and day of week to reclassify them into discretionary trips if necessary.

The methodology to automatically identify trip purpose was applied to the data collected by the University of Minnesota travel behavior study on the use of the I-35W Bridge as a case study. The results from the methodology were consistent with 66.7% of the trip purpose from the travel diary data. Detailed analysis of the trips with travel diary trip purpose different from the automatically calculated trip purpose suggests that travel diary trip purpose data are not always the ground truth. The analysis indicates that 41% or more of the trips for which the trip purpose data did not match between the travel diary and the methodology may be misreported in the travel diary. In traditional demand modeling, the aggregate number of trips by purpose are used and the errors due to potentially misreported trips may cancel out if the distribution of the trip purpose is close to reality. The methodology was limited by its ability to discern social, or pickup drop-off components of a trip if it ended in locations with other types of activities. The underlying quality of the spatial data in the commercial mapping software also affected the results from the methodology. Some changes were made to the methodology before it was applied to the Commute Atlanta data, to reflect the availability of habitual behavior information available in the Commute Atlanta data.

The dissertation explored methods to quantify intra-household travel behavior variability designed to represent the variability-seeking nature of the driver and travel constraints related to the built environment around a household. A literature review of variability estimation methods identified previous uses of mean absolute deviation, standard deviation and coefficient of variation to measure the variability in number of trips, with daily distance traveled and total travel time as the most suitable methods. Nonparametric methods such as bootstrap, Mann Whitney U test, and Spearman's Coefficient were found to be the most appropriate to assess intra-household variability with respect to socio-economic variables.

The intra-household travel behavior variability in the Commute Atlanta data was then analyzed. The results of the daily distance traveled are influenced by the dispersion of the activity locations, and the daily vehicle hours traveled is influenced by the road characteristics, congestion, and distance traveled. The variability of the number of trips is the estimate least likely to be influenced by factors other than the variability-seeking nature of the household. The analysis found that values of mean absolute deviation and standard deviation in distance traveled appear to correlate with the mean number of trips of a household. Households that have the same numerical values for these variability measures may not have the same amount of variability in their travel. The analyses indicate that the coefficient of variation may be the most suitable measure of dispersion in travel behavior to compare travel behavior variability across two households. Comparison of the household demographics and the coefficient of variation of the number of trips per day found correlation between household income, household size, household vehicle ownership, and the presence or absence of children and workers.

The dissertation then explored activity space methods to answer the ‘Where’ part of the ‘When’, ‘Where’, ‘How’ and ‘Why’ questions that planners and modelers are trying to answer about travel behavior. The standard Confidence Ellipse method is a good method for illustrating the spatial dispersion of activity locations, but the ellipse includes areas that are never visited by the household and the ellipse is not very useful in assessing the changes in activity frequency. The Kernel Density method was better for assessing the activity frequency and the number of locations, but did not capture the spatial dispersion of activities. The dissertation developed a new methodology, Modified Kernel Density, which includes the Kernel Density area of trip-end activities as well as the Kernel Density area of the observed travel activity. The method considers the travel between activities as an activity in itself and includes the travel space in the overall activity space estimation.

A regression tree analysis of the activity space from the above three methods when applied to the Commute Atlanta data found that vehicle ownership was the most significant factor affecting activity space, followed by income, household size and number of workers. A correlation analysis found that Confidence Ellipse method poorly correlates with the number of trips and only does marginally better for daily distance traveled. The Kernel Density correlated well with the number of trips but was not as good for daily distance traveled. The Modified Kernel Density area correlated well with both the number of trips per day and the daily distance traveled. The above results were also observed in a case study that found the changes in Modified Kernel Density area being similar to the changes in the number of trips and daily distance traveled for a sample household.

The dissertation next developed activity participation models using structural equation modeling, culminating with models that include travel variability and activity space as explanatory variables. The five steps used in developing structural equation models are model specification, model identification, model estimation, model testing and model modification. The dissertation identified exogenous, endogenous and latent variables from existing research, and the correlations found in this dissertation work. The variability-seeking nature was presented as a latent variable in this modeling effort. The dissertation developed five activity participation models that reflect the activity based travel demand modeling theory and used AMOS software to estimate the models.

The first model incorporated socio-demographic variables and week day as exogenous variables, the activity participation durations and the resulting number of trips as endogenous variables. The model's goodness-of-fit was acceptable while the RMSEA was above acceptable limits.

The second model incorporated the travel times to maintenance and discretionary activities along with the activity participation durations as endogenous variables to reflect the role of the travel time durations on the activity participation durations. This model had acceptable goodness-of-fit and RMSEA values. The model verified the hypothesis that when a household can allocate more time to travel per day the individual is able to make more trips per day.

The third model incorporated the variability-seeking nature of the household as a latent variable indicated by the coefficient of variation of trips per day and coefficient of variation of daily distance traveled. The third model yielded a better goodness-of-fit than the second model. The impact of the variability-seeking nature on the coefficient of variation of trips per day and the coefficient of variation of the daily distance traveled was quantified and it appears to be significant.

The fourth model incorporated the activity space measures to the second model as endogenous variables that are impacted by demographics and which in turn impact the activity participation durations. The fourth model had a better goodness-of-fit and RMSEA values than the third model. The Confidence Ellipse activity space appears to have significant impact on maintenance and discretionary activity durations but no significant impact on repetitious activities such as home, work and school activity. The Kernel Density and Modified Kernel Density did not appear to significantly impact home and maintenance activity durations. The Kernel Density appears to have significant impact on the work activity, while the Modified Kernel Density appears to have significant impact on discretionary, and school activity durations as well as the travel durations. The Modified Kernel Density appears to be better at explaining activity participation than the Kernel Density method.

The fifth model incorporates both the activity space and variability-seeking nature variables in the model. The model had the best goodness-of-fit and RMSEA value. The parameter estimates from the fifth model were consistent with results from the third and fourth model. Comparing the five models based on goodness-of-fit and model parsimony the dissertation found that that fourth model with activity space incorporated in the activity participation model improved the modeling process while not over fitting. The dissertation recommends the use of activity space to explain the spatial appetite of the households in the activity participation modeling process.

The final recommended activity participation model included Modified Kernel Density area to represent the spatial appetite. This model was generated based on the model with activity space after eliminating non-significant relationships, Confidence Ellipse area and Kernel Density area variables. The recommended model had the best RMSEA, AGFI, ACI, and Hoeler's critical N which suggest that this model is the most parsimonious model that best fits the data.

Research Contributions

The contributions of this dissertation are in the areas of processing of longitudinal travel data, trip purpose identification, estimating travel behavior variability, estimating activity space and incorporating the spatial appetite and variability-seeking nature of households in the activity participation modeling process.

The dissertation contributes to the existing methodologies for the processing of passively collected longitudinal travel data into a more useful format. Passively collected travel data usually include details such as the GPS coordinates, vehicle speed, GPS characteristics, and other data from the on-board diagnostics computer. The travel data also has household demographics collected at the time of recruitment and through follow up surveys. The data need to be processed before they can be used in travel behavior studies. The dissertation details methods that automate data processing, identifying trip chains, identifying the vehicle and household, joining socio-demographic data to the travel data and summarizing the information in a useful format for model development.

A methodology to automatically identify likely activities at the end of trips using passively collected longitudinal data and commercially available mapping data is another contribution by this dissertation. When passively collecting instrumented vehicle data, information about the activity undertaken at the end of the trip is usually not collected. The data are of little value to modeling and planning purpose without identifying the activities at the trip ends. Existing methods using regional geographically referenced data are affected by the quality of the data, and the inconsistencies of the data formats across regions. The existing methods are not easily transferable across different parts of the United States and the results are not comparable across regions. The dissertation created a methodology that utilizes information from commercial software that can be replicated elsewhere. The results from using this methodology found that individuals could potentially report travel purposes incorrectly even if the travel data are presented in a map. The analysis of the trip purpose estimated by the new methodology underlines the importance of using automated tools to double check travel diary entries when the participant is completing the travel diaries.

The dissertation explored methods that help estimate different travel behavior variability measures and studied their potential in explaining the variability-seeking nature of a household. While standard measures such as mean absolute deviation and standard deviation provide a good estimate of the variability they are biased by the underlying mean number of trips per household.

The dissertation found that the coefficient of variation of number of trips per day and coefficient of variation of daily distance appear to be useful measures of travel behavior variability.

The dissertation developed the Modified Kernel Density method to estimate activity space, so as to capture the entire activity extent of a household. Existing methods, such as Confidence Ellipse and Kernel Density, suffer from potential issues in activity space estimation. The Confidence Ellipse method captures the spatial dispersion of the activity locations but does not accurately capture the activity frequency and number of locations. Ellipses also include large areas that households never visit. The Kernel Density method captures the number of activity locations and activity frequency, but does not capture accurately capture the spatial dispersion of the activities. The new method integrates the travel activity and the activity locations as part of the activity space by giving different bandwidths to the influence of each location. The results from this methodology appear to better capture the changes in household travel behavior than the existing standard methods. The use of high-resolution GPS data provide better activity space estimates than just using the Origin and Destination information collected by traditional travel diaries and should be considered when designing travel diary data collection systems.

A methodology to create activity participation model that incorporates variables to explain the travel variability-seeking nature and spatial appetite of households was developed by this dissertation. Current activity participation models do not reflect the variability-seeking nature nor spatial appetite of the individual household which is then left as part of the error term. The dissertation hypothesized that the measure of activity space and measures of temporal variations may represent the spatial appetite and variability-seeking nature of the individual household and using the above variables in the modeling process may help in building travel behavior models that may capture activity participation more accurately. The results from the activity participation models developed in this dissertation suggest that including Modified Kernel Density area as activity space in the modeling process may improve the model fit, while not over-fitting.

With respect to study limitations, the travel data used in this dissertation are only from personal vehicles and the study does not capture travel by other modes. The results from this dissertation are a case study and exploratory in nature since the sample is small and is not representative of Atlanta. The author also acknowledges that the data from Commute Atlanta study is large longitudinally but with small number of households. Hence the results from the model may not be reproduced if data from larger number of households are used in the modeling process. However the methodologies provide an excellent framework for processing passively collected GPS data, converting them to a useful format for modeling, estimate travel behavior variability and activity space measures, and incorporate preferences of the individual driver such as spatial appetite and variability-seeking nature into the modeling process. Using the methodologies developed in this dissertation with larger, and more comprehensive datasets should provide valuable insight into the impacts of the personal preferences of individuals on their activity choices and travel behavior. The assumptions used in this dissertation such as the vehicle travel representing the complete travel of the participant, and activity location being within a quarter mile of parking need to be re-evaluated and modified if data from different sources or regions are used. Therefore, adequate care must be taken with the underlying assumptions before using the methodologies developed in this dissertation.

Recommendations for Future Work

The dissertation had to make assumptions to address the limitation of the data and information available. This section makes recommendations for future work to improve understanding of travel behavior and to create better activity participation models.

One of the limitations of using data from instrumented vehicle studies is missing travel data from other modes. With improvements in smartphone technology and the increasing market penetration of the smartphones, it is possible to capture travel data across all modes by using smartphone apps such as the Commute Warrior [104]. Assumptions employed in the research methodologies should be re-evaluated and modified as necessary based on future research findings. The results from using data across all modes will help in better understanding the activity participation behavior. The mode for a particular tour is likely to significantly impact the activities and the parameters used in modeling activities will differ across the modes.

The dissertation research found that a significant number of trips have the travel diary data misreported, even when the travel activity was presented in a map. There is a potential for missing GPS data that may have impacted some of these trips. Future studies should develop methodologies to pre-process the spatial data around trip ends and offer information about the businesses and other land uses in the neighborhood of the trip end to the user while they are completing their travel diaries. This may help in reminding the user of what activities they undertook and help in accurately capturing travel dairy data.

The analyses found that identifying pickup and drop-off activities are difficult to automate, given that the spatial data around the trip end often do not provide any clues especially in residential and mixed use areas. Hence, a detailed study of pickup drop-off activities and their impacts on travel behavior modeling process. Collecting travel data from all participants in a household will likely help in identifying pickup and drop off activities, as will collection of electronic travel diary trip purpose data.

Future research is needed in developing methodologies for probability-based models for trip purpose identification. A probability-based model, such as one using the Bayesian probability, may be able to assign the probability of an activity occurring at a location given the locational characteristics, time of day, day of week and other socio-demographic information of the household. The model could also utilize the business hours of operation to more accurately identify trip purpose. The probability based trip purpose methods will explicitly address the uncertainty in the automatic trip purpose identification methods and provide modelers insight into variability of the trip purpose to the same location by different households and at different time of day.

Future work should explore methods to capture variability in route and mode choices. The variation in route and mode choice may explain the impact of congestion, infrastructure availability, and built environment on the travel behavior of individual households. Incorporating the variability in route and mode choice may help in better understanding of the activity participation decisions made by households and therefore help in the formulation of better policy and planning.

While the analyses used activity space to represent the spatial appetite of the individual households, the impact of spatial appetite (which is a latent variable) could not be measured in the modeling process due to the lack of other indicator variables. The dissertation recommends evaluating other variables such as variability in the number of locations and number of unique locations visited by a household that indicate the spatial appetite of a household and use in the activity participation modeling process.

Travel mode and availability of travel modes play a critical role in trip chaining, available time for activity participation, and the difficulty of accessing activities. Incorporating travel mode choices from longitudinal travel data into the activity participation modeling process and evaluating their impact needs to be further studied.

In the last decade, travel demand modeling research and applications have increasingly moved towards behavior-based modeling techniques from traditional trip based modeling. However, most of the activity-based models use socio-demographic variables to explain activity participation and trip generation. Behavioral variability of the individuals and households have not been explicitly incorporated into the modeling process. The behavioral variability of the individuals may have as much impact on their travel choices as socio-demographic and seasonal elements. The dissertation research has laid out a framework to estimate the behavioral preferences of the individual households and used them in the modeling process. The next steps would be to incorporate more behavioral characteristics, such as route choice, mode choice, joint activity choices, etc., as variables into the activity participation and trip generation modeling process and studying their impact on travel behavior. The results from travel demand models that incorporate behavioral variables may provide planners and policy makers' insights into the effectiveness of different policies on the changes in travel behavior.

APPENDIX A

ACTIVITY PARTICIPATION MODELING DETAILED TABLES

Table A.1 Estimates of base Activity Participation Model

			Estimate	S.E.	C.R.	P
Home Activity	<---	Number of Vehicles	-36.800	2.318	-15.879	***
Home Activity	<---	Household Size	31.972	2.871	11.138	***
Home Activity	<---	Number of Children	.715	3.665	.195	.845
Home Activity	<---	Low Income <30K	34.387	6.651	5.171	***
Home Activity	<---	Middle Income $30K to $75K	-45.737	4.145	-11.034	***
Home Activity	<---	High Income $75K to $100K	28.506	5.256	5.424	***
Home Activity	<---	Work Status	-100.774	4.064	-24.794	***
Home Activity	<---	Student Status	33.974	6.106	5.564	***
Home Activity	<---	College Educated	36.332	3.347	10.853	***
Home Activity	<---	Female	31.853	3.380	9.425	***
Home Activity	<---	Young <35 years	-189.589	4.797	-39.523	***
Home Activity	<---	Middle Age 35 to 55 years	-165.705	4.053	-40.882	***
Home Activity	<---	Weekday	-128.695	3.695	-34.832	***
Work Activity	<---	Number of Vehicles	4.359	1.269	3.434	***
Work Activity	<---	Household Size	-11.362	1.573	-7.221	***
Work Activity	<---	Number of Children	-5.323	2.010	-2.648	.008
Work Activity	<---	Low Income <30K	3.169	3.641	.870	.384
Work Activity	<---	Middle Income $30K to $75K	24.327	2.269	10.721	***
Work Activity	<---	High Income $75K to $100K	-12.282	2.877	-4.269	***
Work Activity	<---	Work Status	110.422	2.233	49.461	***
Work Activity	<---	Student Status	-29.467	3.385	-8.705	***
Work Activity	<---	College Educated	4.232	1.832	2.310	.021
Work Activity	<---	Female	-3.726	1.871	-1.991	.046
Work Activity	<---	Young <35 years	84.101	2.970	28.321	***
Work Activity	<---	Middle Age 35 to 55 years	93.079	2.273	40.951	***
Work Activity	<---	Weekday	138.858	2.022	68.660	***
Maintenance Activity	<---	Weekday	.260	2.041	.127	.899
Maintenance Activity	<---	Middle Age 35 to 55 years	5.702	2.031	2.808	.005
Maintenance Activity	<---	Female	13.794	1.892	7.292	***
Maintenance Activity	<---	College Educated	-18.369	1.849	-9.934	***
Maintenance Activity	<---	Student Status	-14.846	3.213	-4.621	***
Maintenance Activity	<---	Work Status	-23.595	2.217	-10.643	***
Maintenance Activity	<---	High Income $75K to $100K	12.232	2.903	4.213	***
Maintenance Activity	<---	Middle Income $30K to $75K	2.005	2.289	.876	.381
Maintenance Activity	<---	Low Income <30K	-29.670	3.674	-8.077	***
Maintenance Activity	<---	Number of Children	-10.946	2.011	-5.443	***
Maintenance Activity	<---	Household Size	7.654	1.578	4.850	***
Maintenance Activity	<---	Number of Vehicles	-7.333	1.277	-5.742	***
Discretionary Activity	<---	Number of Vehicles	38.163	1.493	25.564	***
Discretionary Activity	<---	Household Size	-30.515	1.851	-16.490	***

Table A.1 continued

			Estimate	S.E.	C.R.	P
Discretionary Activity	<---	Number of Children	14.548	2.364	6.154	***
Discretionary Activity	<---	Low Income <**3T0K**	21.910	4.281	5.118	***
Discretionary Activity	<---	Middle Income $30K to $75K	22.875	2.668	8.573	***
Discretionary Activity	<---	High Income $75K to $100K	-26.608	3.383	-7.865	***
Discretionary Activity	<---	Work Status	-1.390	2.626	-.529	.597
Discretionary Activity	<---	Student Status	-25.410	3.986	-6.375	***
Discretionary Activity	<---	College Educated	-30.643	2.155	-14.221	***
Discretionary Activity	<---	Female	-41.538	2.204	-18.843	***
Discretionary Activity	<---	Young <35 years	112.653	3.530	31.917	***
Discretionary Activity	<---	Middle Age 35 to 55 years	67.411	2.679	25.162	***
Discretionary Activity	<---	Weekday	-29.367	2.378	-12.349	***
School Activity	<---	Weekday	2.515	.795	3.165	.002
School Activity	<---	Middle Age 35 to 55 years	-1.804	.895	-2.015	.044
School Activity	<---	Young <35 years	-7.985	1.179	-6.771	***
School Activity	<---	College Educated	7.621	.720	10.584	***
School Activity	<---	Student Status	19.830	1.332	14.888	***
School Activity	<---	Work Status	-.307	.878	-.350	.726
School Activity	<---	High Income $75K to $100K	3.544	1.131	3.135	.002
School Activity	<---	Middle Income $30K to $75K	8.488	.892	9.520	***
School Activity	<---	Low Income <30K	.912	1.431	.638	.524
School Activity	<---	Number of Children	.389	.790	.493	.622
School Activity	<---	Household Size	3.271	.618	5.290	***
School Activity	<---	Number of Vehicles	1.898	.499	3.804	***
Number of Trips	<---	Home Activity	-.009	.000	-84.193	***
Number of Trips	<---	Work Activity	-.009	.000	-66.829	***
Number of Trips	<---	Maintenance Activity	-.008	.000	-57.115	***
Number of Trips	<---	Discretionary Activity	-.009	.000	-71.887	***
Number of Trips	<---	School Activity	-.008	.000	-37.549	***

Table A.2 Total Effects of base Activity Participation Model

	Weekday	Middle Age 35 to 55 years	Young <35 years	Female	College Educated	Student Status	Work Status	High Income $75K to $100K	Middle Income $30K to $75K	Low Income <30K	Number of Vehicles	Number of Children	Household Size	School	Discretionary	Maintenance	Work	Home
School	2.5	-1.8	-8.0	0.0	7.6	19.8	-0.3	3.5	8.5	0.9	1.9	0.4	3.3	0.0	0.0	0.0	0.0	0.0
Discretionary	-29.4	67.4	112.7	-41.5	-30.6	-25.4	-1.4	-26.6	22.9	21.9	38.2	14.5	-30.5	0.0	0.0	0.0	0.0	0.0
Maintenance	0.3	5.7	0.0	13.8	-18.4	-14.8	-23.6	12.2	2.0	-29.7	-7.3	-10.9	7.7	0.0	0.0	0.0	0.0	0.0
Work	138.9	93.1	84.1	-3.7	4.2	-29.5	110.4	-12.3	24.3	3.2	4.4	-5.3	-11.4	0.0	0.0	0.0	0.0	0.0
Home	-128.7	-165.7	-189.6	31.9	36.3	34.0	-100.8	28.5	-45.7	34.4	-36.8	0.7	32.0	0.0	0.0	0.0	0.0	0.0
Number of Trips	0.263	0.112	0.09	0.008	-0.017	0.126	0.192	-0.039	-0.068	-0.331	0.001	-0.014	-0.007	-0.008	-0.009	-0.008	-0.009	-0.009

Table A.3 Indirect Effects of base Activity Participation Model

	Weekday	Middle Age 35 to 55 years	Young <35 years	Female	College Educated	Student Status	Work Status	High Income $75K to $100K	Middle Income $30K to $75K	Low Income <30K	Number of Vehicles	Number of Children	Household Size	School	Discretionary	Maintenance	Work	Home
School	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Discretionary	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Maintenance	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Work	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Home	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Number of Trips	0.263	0.112	0.09	0.008	-0.017	0.126	0.192	-0.039	-0.068	-0.331	0.001	-0.014	-0.007	0	0	0	0	0

Table A.4 Estimates of Activity Participation Model with Travel Duration

			Estimate	S.E.	C.R.	P
Home Activity	<---	Number of Vehicles	-36.800	2.318	-15.877	***
Home Activity	<---	Household Size	31.972	2.946	10.853	***
Home Activity	<---	Number of Children	.715	3.684	.194	.846
Home Activity	<---	Low Income <30K	34.387	6.656	5.166	***
Home Activity	<---	Middle Income $30K to $75K	-45.737	4.135	-11.060	***
Home Activity	<---	High Income $75K to $100K	28.506	5.235	5.445	***
Home Activity	<---	Work Status	-100.774	4.052	-24.870	***
Home Activity	<---	Student Status	33.974	6.072	5.595	***
Home Activity	<---	College Educated	36.332	3.346	10.859	***
Home Activity	<---	Female	31.853	3.380	9.425	***
Home Activity	<---	Young <35 years	-189.589	4.974	-38.117	***
Home Activity	<---	Middle Age 35 to 55 years	-165.705	4.340	-38.181	***
Home Activity	<---	Weekday	-128.695	3.695	-34.832	***
Work Activity	<---	Number of Vehicles	4.359	1.269	3.436	***
Work Activity	<---	Household Size	-11.362	1.613	-7.046	***
Work Activity	<---	Number of Children	-5.323	2.021	-2.634	.008
Work Activity	<---	Low Income <30K	3.169	3.644	.870	.384
Work Activity	<---	Middle Income $30K to $75K	24.327	2.264	10.747	***
Work Activity	<---	High Income $75K to $100K	-12.282	2.865	-4.286	***
Work Activity	<---	Work Status	110.422	2.227	49.573	***
Work Activity	<---	Student Status	-29.467	3.362	-8.765	***
Work Activity	<---	College Educated	4.232	1.831	2.311	.021
Work Activity	<---	Female	-3.726	1.871	-1.991	.046
Work Activity	<---	Young <35 years	84.101	3.079	27.314	***
Work Activity	<---	Middle Age 35 to 55 years	93.079	2.449	38.002	***
Work Activity	<---	Weekday	138.858	2.022	68.660	***
Maintenance Activity	<---	Weekday	.260	2.041	.127	.899
Maintenance Activity	<---	Middle Age 35 to 55 years	5.702	2.109	2.703	.007
Maintenance Activity	<---	Female	13.794	1.892	7.292	***
Maintenance Activity	<---	College Educated	-18.369	1.848	-9.939	***
Maintenance Activity	<---	Student Status	-14.846	3.213	-4.621	***
Maintenance Activity	<---	Work Status	-23.595	2.204	-10.706	***
Maintenance Activity	<---	High Income $75K to $100K	12.232	2.892	4.230	***
Maintenance Activity	<---	Middle Income $30K to $75K	2.005	2.284	.878	.380
Maintenance Activity	<---	Low Income <30K	-29.670	3.677	-8.069	***
Maintenance Activity	<---	Number of Children	-10.946	2.020	-5.419	***
Maintenance Activity	<---	Household Size	7.654	1.627	4.703	***
Maintenance Activity	<---	Number of Vehicles	-7.333	1.280	-5.727	***
Discretionary Activity	<---	Number of Vehicles	38.163	1.492	25.580	***
Discretionary Activity	<---	Household Size	-30.515	1.896	-16.093	***
Discretionary Activity	<-	Number of Children	14.548	2.377	6.121	***
Discretionary Activity	<---	Low Income <30K	21.910	4.285	5.114	***
Discretionary Activity	<---	Middle Income $30K to $75K	22.875	2.662	8.594	***
Discretionary Activity	<---	High Income $75K to $100K	-26.608	3.370	-7.897	***
Discretionary Activity	<---	Work Status	-1.390	2.620	-.530	.596

Table A.4 continued

			Estimate	S.E.	C.R.	P
Discretionary Activity	<---	Student Status	-25.410	3.958	-6.420	***
Discretionary Activity	<---	College Educated	-30.643	2.154	-14.228	***
Discretionary Activity	<---	Female	-41.538	2.204	-18.843	***
Discretionary Activity	<---	Young <35 years	112.653	3.660	30.781	***
Discretionary Activity	<---	Middle Age 35 to 55 years	67.411	2.889	23.337	***
Discretionary Activity	<---	Weekday	-29.367	2.378	-12.349	***
School Activity	<---	Weekday	2.515	.795	3.165	.002
School Activity	<---	Middle Age 35 to 55 years	-1.804	.965	-1.869	.062
School Activity	<---	Young <35 years	-7.985	1.223	-6.530	***
School Activity	<---	College Educated	7.621	.720	10.590	***
School Activity	<---	Student Status	19.830	1.323	14.994	***
School Activity	<---	Work Status	-.307	.876	-.351	.726
School Activity	<---	High Income $75K to $100K	3.544	1.126	3.147	.002
School Activity	<---	Middle Income $30K to $75K	8.488	.890	9.543	***
School Activity	<---	Low Income <30K	.912	1.432	.637	.524
School Activity	<---	Number of Children	.389	.794	.490	.624
School Activity	<---	Household Size	3.271	.634	5.162	***
School Activity	<---	Number of Vehicles	1.898	.499	3.807	***
Travel Time to Maintenance	<---	Number of Vehicles	.802	.182	4.410	***
Travel Time to Maintenance	<---	Household Size	-.140	.231	-.605	.545
Travel Time to Maintenance	<---	Number of Children	-.419	.289	-1.446	.148
Travel Time to Maintenance	<---	Low Income <30K	-4.620	.522	-8.851	***
Travel Time to Maintenance	<---	Middle Income $30K to $75K	.645	.324	1.989	.047
Travel Time to Maintenance	<---	High Income $75K to $100K	-1.892	.411	-4.608	***
Travel Time to Maintenance	<---	Work Status	.528	.319	1.655	.098
Travel Time to Maintenance	<---	Student Status	-3.719	.481	-7.736	***
Travel Time to Maintenance	<---	College Educated	-5.229	.262	-19.928	***
Travel Time to Maintenance	<---	Female	-2.270	.269	-8.452	***
Travel Time to Maintenance	<---	Young <35 years	2.560	.433	5.914	***
Travel Time to Maintenance	<---	Middle Age 35 to 55 years	4.387	.349	12.564	***
Travel Time to Maintenance	<---	Weekday	4.399	.290	15.184	***
Travel Time to Discretionary	<---	Number of Vehicles	1.832	.135	13.578	***
Travel Time to Discretionary	<---	Household Size	-1.598	.171	-9.319	***
Travel Time to Discretionary	<---	Number of Children	1.850	.215	8.606	***
Travel Time to Discretionary	<---	Low Income <30K	-1.632	.387	-4.211	***
Travel Time to Discretionary	<---	Middle Income $30K to $75K	-1.488	.241	-6.183	***
Travel Time to Discretionary	<---	High Income $75K to $100K	-2.226	.305	-7.306	***
Travel Time to Discretionary	<---	Work Status	.663	.237	2.797	.005
Travel Time to Discretionary	<---	Student Status	-.342	.358	-.955	.339
Travel Time to Discretionary	<---	College Educated	-2.751	.195	-14.127	***
Travel Time to Discretionary	<---	Female	-1.575	.199	-7.898	***
Travel Time to Discretionary	<---	Young <35 years	3.085	.331	9.327	***
Travel Time to Discretionary	<---	Middle Age 35 to 55 years	1.896	.261	7.258	***
Travel Time to Discretionary	<---	Weekday	.022	.215	.104	.917
Number of Trips	<---	Home Activity	-.003	.000	-24.589	***
Number of Trips	<---	Work Activity	-.002	.000	-14.730	***

Table A.4 continued

			Estimate	S.E.	C.R.	P
Number of Trips	<---	Maintenance Activity	-.003	.000	-21.448	***
Number of Trips	<---	Discretionary Activity	-.004	.000	-31.539	***
Number of Trips	<---	School Activity	-.001	.000	-7.041	***
Number of Trips	<---	Travel Time to Maintenance	.053	.000	114.626	***
Number of Trips	<---	Travel Time to Discretionary	.029	.001	46.800	***

Table A.5 Total Effects of Activity Participation Model with Travel Duration

	Weekday	Middle Age 35 to 55 years	Young <35 years	Female	College Educated	Student Status	Work Status	High Income $75K to $100K	Middle Income $30K to $75K	Low Income <30K	Number of Vehicles	Number of Children	Household Size	Travel Discretionary	Travel Maintenance	School	Discretionary	Maintenance	Work	Home
Travel Discretionary	0.0	1.9	3.1	-1.6	-2.8	-0.3	0.7	-2.2	-1.5	-1.6	1.8	1.9	-1.6	0	0	0	0	0	0	0
Travel Maintenance	4.4	4.4	2.6	-2.3	-5.2	-3.7	0.5	-1.9	0.6	-4.6	0.8	-0.4	-0.1	0	0	0	0	0	0	0
School	2.5	-1.8	-8.0	0.0	7.6	19.8	-0.3	3.5	8.5	0.9	1.9	0.4	3.3	0	0	0	0	0	0	0
Discretionary	-29.4	67.4	112.7	-41.5	-30.6	-25.4	-1.4	-26.6	22.9	21.9	38.2	14.5	-30.5	0	0	0	0	0	0	0
Maintenance	0.3	5.7	0.0	13.8	-18.4	-14.8	-23.6	12.2	2.0	-29.7	-7.3	-10.9	7.7	0	0	0	0	0	0	0
Work	138.9	93.1	84.1	-3.7	4.2	-29.5	110.4	-12.3	24.3	3.2	4.4	-5.3	-11.4	0	0	0	0	0	0	0
Home	-128.7	-165.7	-189.6	31.9	36.3	34.0	-100.8	28.5	-45.7	34.4	-36.8	0.7	32.0	0	0	0	0	0	0	0
Number of Trips	0.433	0.302	0.174	-0.127	-0.311	-0.138	0.185	-0.157	-0.031	-0.396	0.062	0.013	-0.03	0.029	0.053	-0.001	-0.004	-0.003	-0.002	-0.003

Table A.6 Indirect Effects of Activity Participation Model with Travel Duration

	Weekday	Middle Age 35 to 55 years	Young <35 years	Female	College Educated	Student Status	Work Status	High Income $75K to $100K	Middle Income $30K to $75K	Low Income <30K	Number of Vehicles	Number of Children	Household Size	Travel Discretionary	Travel Maintenance	School	Discretionary	Maintenance	Work	Home
Travel Discretionary	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Travel Maintenance	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
School	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Discretionary	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Maintenance	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Work	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Home	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Number of Trips	0.433	0.302	0.174	-0.127	-0.311	-0.138	0.185	-0.157	-0.031	-0.396	0.062	0.013	-0.03	0	0	0	0	0	0	0

Table A.7 Estimates of Activity Participation Model with Variability-seeking Nature

			Estimate	S.E.	C.R.	P
Coefficient of Variation Trips	<---	VariabilitySeekingNature	.185	.001	205.719	***
Coefficient of Variation Distance	<---	VariabilitySeekingNature	.185	.001	205.719	***
Coefficient of Variation Trips	<---	Number of Vehicles	-.027	.001	-24.458	***
Coefficient of Variation Trips	<---	Household Size	-.069	.001	-50.378	***
Coefficient of Variation Trips	<---	Number of Children	.069	.002	39.781	***
Coefficient of Variation Trips	<---	Low Income <30K	.134	.003	43.030	***
Coefficient of Variation Trips	<---	Middle Income $30K to $75K	.052	.002	26.902	***
Coefficient of Variation Trips	<---	High Income $75K to $100K	.016	.002	6.440	***
Coefficient of Variation Trips	<---	Work Status	-.055	.002	-28.936	***
Coefficient of Variation Trips	<---	Student Status	.002	.003	.751	.453
Coefficient of Variation Trips	<---	College Educated	-.011	.002	-7.154	***
Coefficient of Variation Trips	<---	Female	.020	.002	12.311	***
Coefficient of Variation Trips	<---	Young <35 years	-.039	.003	-14.747	***
Coefficient of Variation Trips	<---	Middle Age 35 to 55 years	-.004	.002	-2.061	.039
Coefficient of Variation Distance	<---	Number of Vehicles	-.069	.002	-31.338	***
Coefficient of Variation Distance	<---	Household Size	-.058	.003	-20.685	***
Coefficient of Variation Distance	<---	Number of Children	.067	.004	19.256	***
Coefficient of Variation Distance	<---	Low Income <30K	.079	.006	12.497	***
Coefficient of Variation Distance	<---	Middle Income $30K to $75K	-.047	.004	-11.961	***
Coefficient of Variation Distance	<---	High Income $75K to $100K	.066	.005	13.285	***
Coefficient of Variation Distance	<---	Work Status	-.138	.004	-35.733	***
Coefficient of Variation Distance	<---	Student Status	.082	.006	14.128	***
Coefficient of Variation Distance	<---	College Educated	.090	.003	28.283	***
Coefficient of Variation Distance	<---	Female	.017	.003	5.363	***
Coefficient of Variation Distance	<---	Young <35 years	-.187	.005	-34.694	***
Coefficient of Variation Distance	<---	Middle Age 35 to 55 years	-.134	.004	-31.485	***
Home Activity	<---	Number of Vehicles	-35.430	2.339	-15.150	***
Home Activity	<---	Household Size	36.778	3.007	12.232	***
Home Activity	<---	Number of Children	-4.401	3.731	-1.180	.238
Home Activity	<---	Low Income <30K	23.595	6.760	3.490	***
Home Activity	<---	Middle Income $30K to $75K	-50.638	4.185	-12.101	***
Home Activity	<---	High Income $75K to $100K	27.615	5.242	5.268	***
Home Activity	<---	Work Status	-98.671	4.105	-24.038	***
Home Activity	<---	Student Status	33.215	6.080	5.463	***

Table A.7 continued

			Estimate	S.E.	C.R.	P
Home Activity	<---	College Educated	37.931	3.382	11.217	***
Home Activity	<---	Female	30.084	3.384	8.890	***
Home Activity	<---	Young <35 years	-184.054	5.017	-36.688	***
Home Activity	<---	Middle Age 35 to 55 years	-164.983	4.376	-37.702	***
Home Activity	<---	Weekday	-128.634	3.694	-34.822	***
Work Activity	<---	Number of Vehicles	.125	1.274	.098	.922
Work Activity	<---	Household Size	-15.994	1.638	-9.766	***
Work Activity	<---	Number of Children	-.283	2.037	-.139	.889
Work Activity	<---	Low Income <30K	10.173	3.682	2.763	.006
Work Activity	<---	Middle Income $30K to $75K	22.973	2.279	10.079	***
Work Activity	<---	High Income $75K to $100K	-8.506	2.855	-2.980	.003
Work Activity	<---	Work Status	101.751	2.243	45.362	***
Work Activity	<---	Student Status	-25.334	3.351	-7.559	***
Work Activity	<---	College Educated	8.673	1.842	4.709	***
Work Activity	<---	Female	-2.420	1.864	-1.298	.194
Work Activity	<---	Young <35 years	74.142	3.095	23.954	***
Work Activity	<---	Middle Age 35 to 55 years	86.234	2.461	35.045	***
Work Activity	<---	Weekday	138.741	2.012	68.957	***
Maintenance Activity	<---	Weekday	.308	2.039	.151	.880
Maintenance Activity	<---	Middle Age 35 to 55 years	7.190	2.116	3.399	***
Maintenance Activity	<---	Female	13.873	1.892	7.332	***
Maintenance Activity	<---	College Educated	-20.452	1.866	-10.958	***
Maintenance Activity	<---	Student Status	-15.338	3.210	-4.778	***
Maintenance Activity	<---	Work Status	-18.757	2.239	-8.379	***
Maintenance Activity	<---	High Income $75K to $100K	10.811	2.893	3.737	***
Maintenance Activity	<---	Middle Income $30K to $75K	3.371	2.309	1.459	.144
Maintenance Activity	<---	Low Income <30K	-31.116	3.731	-8.340	***
Maintenance Activity	<---	Number of Children	-12.569	2.042	-6.155	***
Maintenance Activity	<---	Household Size	9.571	1.659	5.768	***
Maintenance Activity	<---	Number of Vehicles	-5.377	1.291	-4.167	***
Discretionary Activity	<---	Number of Vehicles	39.781	1.505	26.437	***
Discretionary Activity	<---	Household Size	-28.250	1.935	-14.602	***
Discretionary Activity	<---	Number of Children	12.142	2.407	5.045	***
Discretionary Activity	<---	Low Income <30K	18.064	4.350	4.153	***
Discretionary Activity	<---	Middle Income $30K to $75K	22.628	2.693	8.404	***
Discretionary Activity	<---	High Income $75K to $100K	-27.971	3.373	-8.294	***
Discretionary Activity	<---	Work Status	1.889	2.651	.713	.476
Discretionary Activity	<---	Student Status	-26.775	3.963	-6.756	***
Discretionary Activity	<---	College Educated	-31.870	2.176	-14.648	***
Discretionary Activity	<---	Female	-42.192	2.206	-19.125	***
Discretionary Activity	<---	Young <35 years	116.414	3.694	31.513	***
Discretionary Activity	<---	Middle Age 35 to 55 years	69.647	2.915	23.890	***
Discretionary Activity	<---	Weekday	-29.319	2.377	-12.335	***
School Activity	<---	Weekday	2.509	.794	3.159	.002
School Activity	<---	Middle Age 35 to 55 years	-1.043	.974	-1.071	.284
School Activity	<---	Young <35 years	-7.463	1.234	-6.046	***

Table A.7 continued

			Estimate	S.E.	C.R.	P
School Activity	<---	College Educated	6.909	.727	9.502	***
School Activity	<---	Student Status	19.336	1.324	14.600	***
School Activity	<---	Work Status	-.465	.886	-.525	.600
School Activity	<---	High Income $75K to $100K	3.413	1.127	3.029	.002
School Activity	<---	Middle Income $30K to $75K	9.666	.900	10.743	***
School Activity	<---	Low Income <30K	2.723	1.453	1.873	.061
School Activity	<---	Number of Children	1.171	.804	1.456	.145
School Activity	<---	Household Size	2.409	.646	3.726	***
School Activity	<---	Number of Vehicles	1.831	.503	3.641	***
Travel Time to Maintenance	<---	Number of Vehicles	.524	.182	2.873	.004
Travel Time to Maintenance	<---	Household Size	-1.257	.235	-5.355	***
Travel Time to Maintenance	<---	Number of Children	.673	.292	2.306	.021
Travel Time to Maintenance	<---	Low Income <30K	-2.313	.528	-4.384	***
Travel Time to Maintenance	<---	Middle Income $30K to $75K	1.783	.327	5.461	***
Travel Time to Maintenance	<---	High Income $75K to $100K	-1.792	.409	-4.381	***
Travel Time to Maintenance	<---	Work Status	-.036	.321	-.111	.912
Travel Time to Maintenance	<---	Student Status	-3.883	.479	-8.103	***
Travel Time to Maintenance	<---	College Educated	-5.718	.264	-21.667	***
Travel Time to Maintenance	<---	Female	-1.923	.268	-7.187	***
Travel Time to Maintenance	<---	Young <35 years	2.247	.435	5.168	***
Travel Time to Maintenance	<---	Middle Age 35 to 55 years	4.662	.351	13.296	***
Travel Time to Maintenance	<---	Weekday	4.387	.288	15.217	***
Travel Time to Discretionary	<---	Number of Vehicles	1.781	.136	13.113	***
Travel Time to Discretionary	<---	Household Size	-2.078	.175	-11.898	***
Travel Time to Discretionary	<---	Number of Children	2.298	.217	10.580	***
Travel Time to Discretionary	<---	Low Income <30K	-.606	.393	-1.544	.123
Travel Time to Discretionary	<---	Middle Income $30K to $75K	-.872	.243	-3.585	***
Travel Time to Discretionary	<---	High Income $75K to $100K	-2.265	.304	-7.438	***
Travel Time to Discretionary	<---	Work Status	.546	.239	2.282	.023
Travel Time to Discretionary	<---	Student Status	-.547	.358	-1.530	.126
Travel Time to Discretionary	<---	College Educated	-3.109	.196	-15.826	***
Travel Time to Discretionary	<---	Female	-1.432	.199	-7.189	***
Travel Time to Discretionary	<---	Young <35 years	3.229	.333	9.688	***
Travel Time to Discretionary	<---	Middle Age 35 to 55 years	2.226	.263	8.459	***
Travel Time to Discretionary	<---	Weekday	.018	.215	.086	.932
Home Activity	<---	Coefficient of Variation Trips	80.555	9.784	8.233	***
Work Activity	<---	Coefficient of Variation Trips	-22.306	5.329	-4.186	***
Maintenance Activity	<---	Coefficient of Variation Trips	2.207	5.400	.409	.683
Discretionary Activity	<---	Coefficient of Variation Trips	19.225	6.296	3.054	.002
School Activity	<---	Coefficient of Variation Trips	-16.987	2.102	-8.083	***
Travel Time to Maintenance	<---	Coefficient of Variation Trips	-18.853	.764	-24.692	***
Travel Time to Discretionary	<---	Coefficient of Variation Trips	-9.287	.568	-16.341	***
Home Activity	<---	Coefficient of Variation Distance	-10.076	4.818	-2.091	.036

Table A.7 continued

			Estimate	S.E.	C.R.	P
Work Activity	<---	Coefficient of Variation Distance	-52.663	2.629	-20.035	***
Maintenance Activity	<---	Coefficient of Variation Distance	26.151	2.643	9.896	***
Discretionary Activity	<---	Coefficient of Variation Distance	16.070	3.106	5.174	***
School Activity	<---	Coefficient of Variation Distance	5.797	1.038	5.586	***
Travel Time to Maintenance	<---	Coefficient of Variation Distance	3.180	.376	8.447	***
Travel Time to Discretionary	<---	Coefficient of Variation Distance	2.830	.280	10.095	***
Number of Trips	<---	Home Activity	-.003	.000	-24.589	***
Number of Trips	<---	Work Activity	-.002	.000	-14.730	***
Number of Trips	<---	Maintenance Activity	-.003	.000	-21.446	***
Number of Trips	<---	Discretionary Activity	-.004	.000	-31.540	***
Number of Trips	<---	School Activity	-.001	.000	-7.041	***
Number of Trips	<---	Travel Time to Maintenance	.053	.000	114.624	***
Number of Trips	<---	Travel Time to Discretionary	.029	.001	46.800	***

Table A.8 Total Effects of Activity Participation Model with Variability-seeking Nature

	VariabilitySeekingNature	Weekday	Middle Age 35 to 55 years	Young <35 years	Female	College Educated	Student Status	Work Status	High Income $75K to $100K	Middle Income $30K to $75K	Low Income <30K	Number of Vehicles	Number of Children	Household Size	Coeff. Var. Distance	Coeff. Var. Trips	Travel Discretionary	Travel Maintenance	School	Discretionary	Maintenance	Work	Home
Coeff. Var. Distance	0.19	0.00	-0.13	-0.19	0.02	0.09	0.08	-0.14	0.07	-0.05	0.08	-0.07	0.07	-0.06	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Coeff. Var. Trips	0.19	0.00	0.00	-0.04	0.02	-0.01	0.00	-0.06	0.02	0.05	0.13	-0.03	0.07	-0.07	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Travel Discretionary	-1.20	0.02	1.89	3.06	-1.57	-2.75	-0.33	0.67	-2.23	-1.49	-1.63	1.83	1.85	-1.60	2.83	-9.29	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Travel Maintenance	-2.91	4.39	4.32	2.39	-2.24	-5.22	-3.66	0.57	-1.88	0.65	-4.59	0.81	-0.41	-0.13	3.18	-18.85	0.00	0.00	0.00	0.00	0.00	0.00	0.00
School	-2.08	2.51	-1.75	-7.88	-0.23	7.62	19.78	-0.33	3.53	8.51	0.90	1.88	0.39	3.25	5.80	-16.99	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Discretionary	6.55	-29.32	67.41	112.65	-41.53	-30.64	-25.41	-1.39	-26.61	22.88	21.91	38.16	14.55	-30.52	16.07	19.23	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Maintenance	5.26	0.31	3.68	-4.98	14.37	-18.13	-13.18	-22.49	12.57	2.26	-28.76	-7.24	-10.65	7.91	26.15	2.21	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Work	-13.90	138.74	93.39	84.88	-3.78	4.19	-29.72	110.25	-12.33	24.28	3.03	4.35	-5.37	-11.40	-52.66	-22.31	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Home	13.07	-128.63	-163.98	-185.33	31.50	36.12	32.56	-101.72	28.22	-45.97	33.60	-36.87	0.46	31.77	-10.08	80.56	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Number of Trips	-0.23	0.43	0.30	0.17	-0.13	-0.31	-0.14	0.19	-0.16	-0.03	-0.39	0.06	0.01	-0.03	0.24	-1.50	0.03	0.05	0.00	0.00	0.00	0.00	0.00

Table A.9 Indirect Effects of Activity Participation Model with Variability-seeking Nature

	VariabilitySeekingNature	Weekday	Middle Age 35 to 55 years	Young <35 years	Female	College Educated	Student Status	Work Status	High Income $75K to $100K	Middle Income $30K to $75K	Low Income <30K	Number of Vehicles	Number of Children	Household Size	Coeff. Var. Distance	Coeff. Var. Trips	Travel Discretionary	Travel Maintenance	School	Discretionary	Maintenance	Work	Home
Coeff. Var. Distance	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Coeff. Var. Trips	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Travel Discretionary	-1.20	0.00	-0.34	-0.17	-0.13	0.36	0.21	0.12	0.04	-0.62	-1.02	0.05	-0.45	0.48	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Travel Maintenance	-2.91	0.00	-0.35	0.15	-0.32	0.50	0.22	0.60	-0.09	-1.13	-2.28	0.28	-1.08	1.13	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
School	-2.08	0.00	-0.70	-0.42	-0.23	0.71	0.44	0.14	0.11	-1.16	-1.82	0.05	-0.78	0.85	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Discretionary	6.55	0.00	-2.24	-3.76	0.66	1.23	1.37	-3.28	1.36	0.25	3.85	-1.62	2.41	-2.27	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Maintenance	5.26	0.00	-3.52	-4.98	0.50	2.32	2.16	-3.73	1.76	-1.11	2.36	-1.86	1.92	-1.67	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Work	-13.90	0.00	7.16	10.73	-1.36	-4.48	-4.39	8.50	-3.83	1.31	-7.15	4.22	-5.09	4.60	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Home	13.07	0.00	1.00	-1.28	1.41	-1.81	-0.66	-3.05	0.61	4.67	10.01	-1.44	4.86	-5.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Number of Trips	-0.23	0.43	0.30	0.17	-0.13	-0.31	-0.14	0.19	-0.16	-0.03	-0.39	0.06	0.01	-0.03	0.24	-1.50	0.00	0.00	0.00	0.00	0.00	0.00	0.00

Table A.10 Estimates of Activity Participation Model with Activity Space

			Estimate	S.E.	C.R.	P
Confidence Ellipse Area	<---	Number of Vehicles	301.394	5.185	58.126	***
Confidence Ellipse Area	<---	Household Size	-2.114	6.590	-.321	.748
Confidence Ellipse Area	<---	Number of Children	-23.919	8.247	-2.900	.004
Confidence Ellipse Area	<---	Low Income <30K	152.113	14.891	10.215	***
Confidence Ellipse Area	<---	Middle Income $30K to $75K	19.720	9.251	2.132	.033
Confidence Ellipse Area	<---	High Income $75K to $100K	-27.592	11.711	-2.356	.018
Confidence Ellipse Area	<---	Work Status	59.800	9.077	6.588	***
Confidence Ellipse Area	<---	Student Status	-260.063	13.631	-19.078	***
Confidence Ellipse Area	<---	College Educated	-290.442	7.485	-38.803	***
Confidence Ellipse Area	<---	Female	-126.408	7.663	-16.496	***
Confidence Ellipse Area	<---	Young <35 years	281.197	11.589	24.264	***
Confidence Ellipse Area	<---	Middle Age 35 to 55 years	450.081	9.801	45.920	***
Kernel Density Area	<---	Number of Vehicles	.622	.078	7.948	***
Kernel Density Area	<---	Household Size	7.503	.099	75.420	***
Kernel Density Area	<---	Number of Children	-9.208	.124	-74.403	***
Kernel Density Area	<---	Low Income <30K	-12.421	.225	-55.258	***
Kernel Density Area	<---	Middle Income $30K to $75K	-10.071	.140	-72.120	***
Kernel Density Area	<---	High Income $75K to $100K	-5.967	.177	-33.756	***
Kernel Density Area	<---	Work Status	3.551	.135	26.235	***
Kernel Density Area	<---	Student Status	-2.210	.199	-11.105	***
Kernel Density Area	<---	College Educated	-2.028	.113	-17.952	***
Kernel Density Area	<---	Female	1.622	.116	14.020	***
Kernel Density Area	<---	Young <35 years	1.010	.091	11.045	***
Kernel Density Area	<---	Middle Age 35 to 55 years	-3.396	.134	-25.269	***
Modified Kernel Density area (sq Km)	<---	Number of Vehicles	9.549	.288	33.152	***
Modified Kernel Density area (sq Km)	<---	Household Size	24.164	.366	66.007	***
Modified Kernel Density area (sq Km)	<---	Number of Children	-25.058	.454	-55.144	***
Modified Kernel Density area (sq Km)	<---	Low Income <30K	-38.463	.827	-46.500	***
Modified Kernel Density area (sq Km)	<---	Middle Income $30K to $75K	-27.555	.514	-53.623	***
Modified Kernel Density area (sq Km)	<---	High Income $75K to $100K	-13.792	.651	-21.202	***
Modified Kernel Density area (sq Km)	<---	Work Status	17.309	.496	34.912	***
Modified Kernel Density area (sq Km)	<---	Student Status	-1.704	.723	-2.358	.018
Modified Kernel Density area (sq Km)	<---	College Educated	-9.860	.416	-23.718	***
Modified Kernel Density area (sq Km)	<---	Female	3.541	.426	8.320	***
Modified Kernel Density area (sq Km)	<---	Middle Age 35 to 55 years	-7.219	.475	-15.214	***
Home Activity	<---	Number of Vehicles	-28.969	2.396	-12.092	***

Table A.10 continued

			Estimate	S.E.	C.R.	P
Home Activity	<---	Household Size	33.040	3.085	10.710	***
Home Activity	<---	Number of Children	-.451	3.855	-.117	.907
Home Activity	<---	Low Income <30K	36.428	6.842	5.324	***
Home Activity	<---	Middle Income $30K to $75K	-46.028	4.319	-10.656	***
Home Activity	<---	High Income $75K to $100K	27.840	5.291	5.262	***
Home Activity	<---	Work Status	-97.930	4.092	-23.931	***
Home Activity	<---	Student Status	28.936	6.114	4.733	***
Home Activity	<---	College Educated	29.178	3.385	8.619	***
Home Activity	<---	Female	29.244	3.392	8.622	***
Home Activity	<---	Young <35 years	-184.669	5.010	-36.861	***
Home Activity	<---	Middle Age 35 to 55 years	-156.091	4.446	-35.107	***
Home Activity	<---	Weekday	-128.819	3.689	-34.916	***
Work Activity	<---	Number of Vehicles	9.454	1.309	7.224	***
Work Activity	<---	Household Size	-4.993	1.685	-2.963	.003
Work Activity	<---	Number of Children	-10.713	2.110	-5.077	***
Work Activity	<---	Low Income <30K	-6.562	3.737	-1.756	.079
Work Activity	<---	Middle Income $30K to $75K	18.419	2.360	7.806	***
Work Activity	<---	High Income $75K to $100K	-14.454	2.890	-5.001	***
Work Activity	<---	Work Status	116.616	2.245	51.938	***
Work Activity	<---	Student Status	-28.953	3.381	-8.563	***
Work Activity	<---	College Educated	.240	1.849	.130	.897
Work Activity	<---	Female	-3.559	1.874	-1.899	.058
Work Activity	<---	Young <35 years	84.363	3.095	27.253	***
Work Activity	<---	Middle Age 35 to 55 years	93.159	2.505	37.190	***
Work Activity	<---	Weekday	138.814	2.015	68.878	***
Maintenance Activity	<---	Weekday	.286	2.039	.140	.888
Maintenance Activity	<---	Middle Age 35 to 55 years	5.780	2.159	2.677	.007
Maintenance Activity	<---	Female	13.512	1.899	7.113	***
Maintenance Activity	<---	College Educated	-15.945	1.871	-8.522	***
Maintenance Activity	<---	Student Status	-13.479	3.226	-4.178	***
Maintenance Activity	<---	Work Status	-26.343	2.226	-11.836	***
Maintenance Activity	<---	High Income $75K to $100K	14.843	2.924	5.076	***
Maintenance Activity	<---	Middle Income $30K to $75K	6.497	2.387	2.721	.006
Maintenance Activity	<---	Low Income <30K	-24.569	3.782	-6.497	***
Maintenance Activity	<---	Number of Children	-6.680	2.115	-3.159	.002
Maintenance Activity	<---	Household Size	3.874	1.705	2.272	.023
Maintenance Activity	<---	Number of Vehicles	-9.489	1.324	-7.166	***
Discretionary Activity	<---	Number of Vehicles	33.024	1.542	21.423	***
Discretionary Activity	<---	Household Size	-26.278	1.985	-13.239	***
Discretionary Activity	<---	Number of Children	8.886	2.486	3.574	***
Discretionary Activity	<---	Low Income <30K	12.598	4.402	2.862	.004
Discretionary Activity	<---	Middle Income $30K to $75K	16.077	2.779	5.785	***
Discretionary Activity	<---	High Income $75K to $100K	-30.549	3.404	-8.973	***
Discretionary Activity	<---	Work Status	-1.271	2.646	-.481	.631
Discretionary Activity	<---	Student Status	-24.045	3.987	-6.031	***
Discretionary Activity	<---	College Educated	-27.198	2.178	-12.487	***

Table A.10 continued

			Estimate	S.E.	C.R.	P
Discretionary Activity	<---	Female	-38.532	2.211	-17.426	***
Discretionary Activity	<---	Young <35 years	109.909	3.684	29.832	***
Discretionary Activity	<---	Middle Age 35 to 55 years	58.649	2.959	19.820	***
Discretionary Activity	<---	Weekday	-29.284	2.374	-12.336	***
School Activity	<---	Weekday	2.523	.794	3.177	.001
School Activity	<---	Middle Age 35 to 55 years	-2.132	.990	-2.154	.031
School Activity	<---	Young <35 years	-7.507	1.232	-6.092	***
School Activity	<---	College Educated	8.008	.729	10.989	***
School Activity	<---	Student Status	19.013	1.334	14.254	***
School Activity	<---	Work Status	-1.130	.885	-1.276	.202
School Activity	<---	High Income $75K to $100K	2.844	1.139	2.497	.013
School Activity	<---	Middle Income $30K to $75K	7.958	.930	8.560	***
School Activity	<---	Low Income <30K	.920	1.473	.625	.532
School Activity	<---	Number of Children	-.140	.832	-.169	.866
School Activity	<---	Household Size	3.176	.664	4.783	***
School Activity	<---	Number of Vehicles	.939	.516	1.822	.068
Travel Time to Maintenance	<---	Number of Vehicles	-.617	.186	-3.319	***
Travel Time to Maintenance	<---	Household Size	-1.906	.239	-7.966	***
Travel Time to Maintenance	<---	Number of Children	1.262	.299	4.215	***
Travel Time to Maintenance	<---	Low Income <30K	-1.997	.531	-3.764	***
Travel Time to Maintenance	<---	Middle Income $30K to $75K	2.440	.335	7.284	***
Travel Time to Maintenance	<---	High Income $75K to $100K	-1.072	.410	-2.612	.009
Travel Time to Maintenance	<---	Work Status	-1.037	.319	-3.254	.001
Travel Time to Maintenance	<---	Student Status	-3.434	.479	-7.168	***
Travel Time to Maintenance	<---	College Educated	-3.977	.263	-15.148	***
Travel Time to Maintenance	<---	Female	-2.256	.267	-8.463	***
Travel Time to Maintenance	<---	Young <35 years	2.110	.432	4.889	***
Travel Time to Maintenance	<---	Middle Age 35 to 55 years	4.088	.354	11.550	***
Travel Time to Maintenance	<---	Weekday	4.414	.286	15.425	***
Travel Time to Discretionary	<---	Number of Vehicles	.814	.138	5.897	***
Travel Time to Discretionary	<---	Household Size	-2.761	.178	-15.530	***
Travel Time to Discretionary	<---	Number of Children	2.829	.223	12.703	***
Travel Time to Discretionary	<---	Low Income <30K	.125	.394	.317	.751
Travel Time to Discretionary	<---	Middle Income $30K to $75K	-.425	.249	-1.708	.088
Travel Time to Discretionary	<---	High Income $75K to $100K	-1.838	.305	-6.029	***
Travel Time to Discretionary	<---	Work Status	-.492	.237	-2.074	.038
Travel Time to Discretionary	<---	Student Status	-.384	.357	-1.076	.282
Travel Time to Discretionary	<---	College Educated	-1.948	.195	-9.987	***
Travel Time to Discretionary	<---	Female	-1.562	.198	-7.884	***
Travel Time to Discretionary	<---	Young <35 years	2.936	.330	8.902	***
Travel Time to Discretionary	<---	Middle Age 35 to 55 years	1.755	.265	6.624	***
Travel Time to Discretionary	<---	Weekday	.032	.213	.149	.882
Home Activity	<---	Confidence Ellipse Area	-.022	.002	-10.484	***
Work Activity	<---	Confidence Ellipse Area	-.001	.001	-1.068	.286
Maintenance Activity	<---	Confidence Ellipse Area	.004	.001	3.741	***
Discretionary Activity	<---	Confidence Ellipse Area	.014	.001	9.981	***

Table A.10 continued

			Estimate	S.E.	C.R.	P
School Activity	<---	Confidence Ellipse Area	.000	.000	-.807	.420
Travel Time to Maintenance	<---	Confidence Ellipse Area	.001	.000	7.501	***
Travel Time to Discretionary	<---	Confidence Ellipse Area	.000	.000	3.461	***
Home Activity	<---	Kernel Density Area	.283	.270	1.048	.295
Work Activity	<---	Kernel Density Area	.940	.148	6.369	***
Maintenance Activity	<---	Kernel Density Area	.280	.149	1.883	.060
Discretionary Activity	<---	Kernel Density Area	-1.143	.174	-6.571	***
School Activity	<---	Kernel Density Area	-.438	.058	-7.523	***
Travel Time to Maintenance	<---	Kernel Density Area	-.148	.021	-7.046	***
Travel Time to Discretionary	<---	Kernel Density Area	-.184	.016	-11.786	***
Home Activity	<---	Modified Kernel Density area (sq Km)	-.132	.079	-1.671	.095
Work Activity	<---	Modified Kernel Density area (sq Km)	-.558	.043	-12.943	***
Maintenance Activity	<---	Modified Kernel Density area (sq Km)	.067	.043	1.536	.125
Discretionary Activity	<---	Modified Kernel Density area (sq Km)	.181	.051	3.565	***
School Activity	<---	Modified Kernel Density area (sq Km)	.140	.017	8.219	***
Travel Time to Maintenance	<---	Modified Kernel Density area (sq Km)	.120	.006	19.561	***
Travel Time to Discretionary	<---	Modified Kernel Density area (sq Km)	.106	.005	23.244	***
Number of Trips	<---	Home Activity	-.003	.000	-24.590	***
Number of Trips	<---	Work Activity	-.002	.000	-14.731	***
Number of Trips	<---	Maintenance Activity	-.003	.000	-21.450	***
Number of Trips	<---	Discretionary Activity	-.004	.000	-31.545	***
Number of Trips	<---	School Activity	-.001	.000	-7.041	***
Number of Trips	<---	Travel Time to Maintenance	.053	.000	114.622	***
Number of Trips	<---	Travel Time to Discretionary	.029	.001	46.800	***

Table A.11 Total Effects of Activity Participation Model with Activity Space

	Weekday	Middle Age 35 to 55 years	Young <35 years	Female	College Educated	Student Status	Work Status	High Income $75K to $100K	Middle Income $30K to $75K	Low Income <30K	Number of Vehicles	Number of Children	Household Size	Modified Kernel Densitv area	Kernel Density Area	Confidence Ellipse Area	Travel Discretionary	Travel Maintenance	School	Discretionary	Maintenance	Work	Home
Modified Kernel Density area	0.0	-7.2	0.0	3.5	-9.9	-1.7	17.3	-13.8	-27.6	-38.5	9.5	-25.1	24.2	0.0	0.0	0	0	0	0	0	0	0	0
Kernel Density Area	0.0	-3.4	1.0	1.6	-2.0	-2.2	3.6	-6.0	-10.1	-12.4	0.6	-9.2	7.5	0.0	0.0	0	0	0	0	0	0	0	0
Confidence Ellipse Area	0.0	450.1	281.2	-126.4	-290.4	-260.1	59.8	-27.6	19.7	152.1	301.4	-23.9	-2.1	0.0	0.0	0	0	0	0	0	0	0	0
Travel Discretionary	0.0	1.8	2.9	-1.5	-2.7	-0.3	0.7	-2.2	-1.5	-1.6	1.8	1.9	-1.6	0.1	-0.2	0	0	0	0	0	0	0	0
Travel Maintenance	4.4	4.3	2.3	-2.2	-5.2	-3.6	0.6	-1.9	0.7	-4.6	0.8	-0.4	-0.1	0.1	-0.1	0	0	0	0	0	0	0	0
School	2.5	-1.8	-8.1	-0.2	7.6	19.8	-0.3	3.5	8.5	0.9	1.9	0.4	3.3	0.1	-0.4	0	0	0	0	0	0	0	0
Discretionary	-29.3	67.4	112.6	-41.5	-30.6	-25.4	-1.4	-26.6	22.9	21.9	38.2	14.5	-30.5	0.2	-1.1	0	0	0	0	0	0	0	0
Maintenance	0.3	6.3	1.5	13.6	-18.4	-15.4	-23.9	12.1	1.9	-29.9	-7.4	-11.0	7.6	0.1	0.3	0	0	0	0	0	0	0	0
Work	138.8	93.4	85.0	-3.9	4.2	-29.8	110.2	-12.3	24.3	3.0	4.3	-5.4	-11.4	-0.6	0.9	0	0	0	0	0	0	0	0
Home	-128.8	-166.1	-190.7	32.1	36.4	34.3	-100.5	28.6	-45.7	34.6	-36.8	0.8	32.0	-0.1	0.3	0	0	0	0	0	0	0	0
Number of Trips	0.43	0.29	0.15	-0.12	-0.31	-0.13	0.19	-0.16	-0.03	-0.39	0.06	0.01	-0.03	0.01	-0.01	0	0	0	0	0	0	0	0

Table A.12 Indirect Effects of Activity Participation Model with Activity Space

	Weekday	Middle Age 35 to 55 years	Young <35 years	Female	College Educated	Student Status	Work Status	High Income $75K to $100K	Middle Income $30K to $75K	Low Income <30K	Number of Vehicles	Number of Children	Household Size	Modified Kernel Density area	Kernel Density Area	Confidence Ellipse Area	Travel Discretionary	Travel Maintenance	School	Discretionary	Maintenance	Work	Home
Modified Kernel Density area	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Kernel Density Area	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Confidence Ellipse Area	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Travel Discretionary	0	0.05	-0.07	0.02	-0.79	0.12	1.20	-0.37	-1.05	-1.72	1.02	-0.97	1.17	0	0	0	0	0	0	0	0	0	0
Travel Maintenance	0	0.20	0.20	0.03	-1.24	-0.20	1.62	-0.80	-1.78	-2.58	1.42	-1.67	1.78	0	0	0	0	0	0	0	0	0	0
School	0	0.31	-0.55	-0.17	-0.38	0.83	0.84	0.70	0.56	0.01	0.95	0.54	0.09	0	0	0	0	0	0	0	0	0	0
Discretionary	0	8.74	2.70	-2.94	-3.44	-1.34	-0.11	3.95	6.79	9.32	5.14	5.66	-4.23	0	0	0	0	0	0	0	0	0	0
Maintenance	0	0.54	1.51	0.14	-2.50	-1.87	2.41	-2.71	-4.57	-5.38	2.13	-4.36	3.71	0	0	0	0	0	0	0	0	0	0
Work	0	0.27	0.60	-0.29	3.95	-0.81	-6.39	2.11	5.87	9.58	-5.11	5.34	-6.41	0	0	0	0	0	0	0	0	0	0
Home	0	-10.05	-5.99	2.81	7.21	5.40	-2.61	0.74	0.34	-1.85	-7.81	1.23	-1.01	0	0	0	0	0	0	0	0	0	0
Number of Trips	0.434	0.29	0.15	-0.12	-0.31	-0.13	0.19	-0.16	-0.03	-0.39	0.06	0.01	-0.03	0.01	-0.012	0	0	0	0	0	0	0	0

Table A.13 Estimates of Activity Participation Model with Variability-seeking Nature and Activity Space

			Estimate	S.E.	C.R.	P
Confidence Ellipse Area	<---	Number of Vehicles	301.394	5.185	58.126	***
Confidence Ellipse Area	<---	Household Size	-2.114	6.590	-.321	.748
Confidence Ellipse Area	<---	Number of Children	-23.919	8.247	-2.900	.004
Confidence Ellipse Area	<---	Low Income <30K	152.113	14.891	10.215	***
Confidence Ellipse Area	<---	Middle Income $30K to $75K	19.720	9.251	2.132	.033
Confidence Ellipse Area	<---	High Income $75K to $100K	-27.592	11.711	-2.356	.018
Confidence Ellipse Area	<---	Work Status	59.800	9.077	6.588	***
Confidence Ellipse Area	<---	Student Status	-260.063	13.631	-19.078	***
Confidence Ellipse Area	<---	College Educated	-290.442	7.485	-38.803	***
Confidence Ellipse Area	<---	Female	-126.408	7.663	-16.496	***
Confidence Ellipse Area	<---	Young <35 years	281.197	11.589	24.264	***
Confidence Ellipse Area	<---	Middle Age 35 to 55 years	450.081	9.801	45.920	***
Kernel Density Area	<---	Number of Vehicles	.622	.078	7.948	***
Kernel Density Area	<---	Household Size	7.503	.099	75.420	***
Kernel Density Area	<---	Number of Children	-9.208	.124	-74.403	***
Kernel Density Area	<---	Low Income <30K	-12.421	.225	-55.258	***
Kernel Density Area	<---	Middle Income $30K to $75K	-10.071	.140	-72.120	***
Kernel Density Area	<---	High Income $75K to $100K	-5.967	.177	-33.756	***
Kernel Density Area	<---	Work Status	3.551	.135	26.235	***
Kernel Density Area	<---	Student Status	-2.210	.199	-11.105	***
Kernel Density Area	<---	College Educated	-2.028	.113	-17.952	***
Kernel Density Area	<---	Female	1.622	.116	14.020	***
Kernel Density Area	<---	Young <35 years	1.010	.091	11.045	***
Kernel Density Area	<---	Middle Age 35 to 55 years	-3.396	.134	-25.269	***
Modified Kernel Density area (sq Km)	<---	Number of Vehicles	9.549	.288	33.152	***
Modified Kernel Density area (sq Km)	<---	Household Size	24.164	.366	66.007	***
Modified Kernel Density area (sq Km)	<---	Number of Children	-25.058	.454	-55.144	***
Modified Kernel Density area (sq Km)	<---	Low Income <30K	-38.463	.827	-46.500	***
Modified Kernel Density area (sq Km)	<---	Middle Income $30K to $75K	-27.555	.514	-53.623	***
Modified Kernel Density area (sq Km)	<---	High Income $75K to $100K	-13.792	.651	-21.202	***
Modified Kernel Density area (sq Km)	<---	Work Status	17.309	.496	34.912	***
Modified Kernel Density area (sq Km)	<---	Student Status	-1.704	.723	-2.358	.018
Modified Kernel Density area (sq Km)	<---	College Educated	-9.860	.416	-23.718	***
Modified Kernel Density area (sq Km)	<---	Female	3.541	.426	8.320	***

Table A.13 continued

			Estimate	S.E.	C.R.	P
Modified Kernel Density area (sq Km)	<---	Middle Age 35 to 55 years	-7.219	.475	-15.214	***
Coefficient of Variation Trips	<---	Number of Vehicles	-.027	.001	-24.501	***
Coefficient of Variation Trips	<---	Household Size	-.070	.001	-50.468	***
Coefficient of Variation Trips	<---	Number of Children	.069	.002	39.713	***
Coefficient of Variation Trips	<---	Low Income <30K	.134	.003	42.884	***
Coefficient of Variation Trips	<---	Middle Income $30K to $75K	.052	.002	26.836	***
Coefficient of Variation Trips	<---	High Income $75K to $100K	.016	.002	6.371	***
Coefficient of Variation Trips	<---	Work Status	-.056	.002	-29.251	***
Coefficient of Variation Trips	<---	Student Status	.001	.003	.463	.644
Coefficient of Variation Trips	<---	College Educated	-.011	.002	-7.231	***
Coefficient of Variation Trips	<---	Female	.019	.002	12.137	***
Coefficient of Variation Trips	<---	Young <35 years	-.037	.003	-14.121	***
Coefficient of Variation Trips	<---	Middle Age 35 to 55 years	-.003	.002	-1.589	.112
Coefficient of Variation Distance	<---	Number of Vehicles	-.069	.002	-31.361	***
Coefficient of Variation Distance	<---	Household Size	-.058	.003	-20.733	***
Coefficient of Variation Distance	<---	Number of Children	.067	.004	19.213	***
Coefficient of Variation Distance	<---	Low Income <30K	.078	.006	12.420	***
Coefficient of Variation Distance	<---	Middle Income $30K to $75K	-.047	.004	-11.996	***
Coefficient of Variation Distance	<---	High Income $75K to $100K	.066	.005	13.249	***
Coefficient of Variation Distance	<---	Work Status	-.139	.004	-35.895	***
Coefficient of Variation Distance	<---	Student Status	.082	.006	13.984	***
Coefficient of Variation Distance	<---	College Educated	.090	.003	28.243	***
Coefficient of Variation Distance	<---	Female	.017	.003	5.271	***
Coefficient of Variation Distance	<---	Young <35 years	-.185	.005	-34.408	***
Coefficient of Variation Distance	<---	Middle Age 35 to 55 years	-.133	.004	-31.283	***
Coefficient of Variation Trips	<---	VariabilitySeekingNature	.185	.001	205.721	***
Coefficient of Variation Distance	<---	VariabilitySeekingNature	.185	.001	205.721	***
Home Activity	<---	Number of Vehicles	-27.781	2.423	-11.466	***
Home Activity	<---	Household Size	36.778	3.113	11.814	***
Home Activity	<---	Number of Children	-4.937	3.881	-1.272	.203
Home Activity	<---	Low Income <30K	27.466	6.897	3.982	***
Home Activity	<---	Middle Income $30K to $75K	-50.413	4.387	-11.492	***
Home Activity	<---	High Income $75K to $100K	27.395	5.289	5.179	***
Home Activity	<---	Work Status	-97.385	4.142	-23.511	***

Table A.13 continued

			Estimate	S.E.	C.R.	P
Home Activity	<---	Student Status	27.132	6.113	4.438	***
Home Activity	<---	College Educated	31.212	3.415	9.139	***
Home Activity	<---	Female	26.771	3.400	7.874	***
Home Activity	<---	Young <35 years	-178.301	5.038	-35.392	***
Home Activity	<---	Middle Age 35 to 55 years	-154.118	4.484	-34.372	***
Home Activity	<---	Weekday	-128.766	3.688	-34.913	***
Work Activity	<---	Number of Vehicles	3.969	1.316	3.016	.003
Work Activity	<---	Household Size	-7.736	1.691	-4.576	***
Work Activity	<---	Number of Children	-9.182	2.112	-4.347	***
Work Activity	<---	Low Income <30K	-2.465	3.746	-.658	.510
Work Activity	<---	Middle Income $30K to $75K	12.634	2.382	5.303	***
Work Activity	<---	High Income $75K to $100K	-13.925	2.872	-4.848	***
Work Activity	<---	Work Status	107.214	2.257	47.493	***
Work Activity	<---	Student Status	-26.999	3.361	-8.032	***
Work Activity	<---	College Educated	3.859	1.855	2.081	.037
Work Activity	<---	Female	-.975	1.867	-.523	.601
Work Activity	<---	Young <35 years	74.953	3.101	24.173	***
Work Activity	<---	Middle Age 35 to 55 years	84.540	2.515	33.616	***
Work Activity	<---	Weekday	138.680	2.003	69.238	***
Maintenance Activity	<---	Weekday	.349	2.036	.171	.864
Maintenance Activity	<---	Middle Age 35 to 55 years	8.479	2.166	3.914	***
Maintenance Activity	<---	Female	13.116	1.901	6.899	***
Maintenance Activity	<---	College Educated	-18.106	1.885	-9.604	***
Maintenance Activity	<---	Student Status	-12.840	3.221	-3.986	***
Maintenance Activity	<---	Work Status	-20.463	2.258	-9.064	***
Maintenance Activity	<---	High Income $75K to $100K	14.934	2.920	5.115	***
Maintenance Activity	<---	Middle Income $30K to $75K	10.729	2.422	4.430	***
Maintenance Activity	<---	Low Income <30K	-24.344	3.808	-6.394	***
Maintenance Activity	<---	Number of Children	-6.378	2.127	-2.999	.003
Maintenance Activity	<---	Household Size	4.823	1.719	2.806	.005
Maintenance Activity	<---	Number of Vehicles	-6.674	1.337	-4.991	***
Discretionary Activity	<---	Number of Vehicles	34.467	1.559	22.107	***
Discretionary Activity	<---	Household Size	-25.509	2.003	-12.735	***
Discretionary Activity	<---	Number of Children	8.404	2.503	3.357	***
Discretionary Activity	<---	Low Income <30K	11.359	4.438	2.560	.010
Discretionary Activity	<---	Middle Income $30K to $75K	17.494	2.823	6.198	***
Discretionary Activity	<---	High Income $75K to $100K	-30.703	3.403	-9.021	***
Discretionary Activity	<---	Work Status	1.148	2.676	.429	.668
Discretionary Activity	<---	Student Status	-24.627	3.987	-6.177	***
Discretionary Activity	<---	College Educated	-28.112	2.198	-12.792	***
Discretionary Activity	<---	Female	-39.236	2.216	-17.704	***
Discretionary Activity	<---	Young <35 years	112.591	3.710	30.350	***
Discretionary Activity	<---	Middle Age 35 to 55 years	60.973	2.988	20.406	***
Discretionary Activity	<---	Weekday	-29.248	2.373	-12.324	***
School Activity	<---	Weekday	2.517	.794	3.171	.002

Table A.13 continued

			Estimate	S.E.	C.R.	P
School Activity	<---	Middle Age 35 to 55 years	-1.868	.999	-1.869	.062
School Activity	<---	Young <35 years	-7.466	1.241	-6.018	***
School Activity	<---	College Educated	7.531	.735	10.245	***
School Activity	<---	Student Status	18.943	1.334	14.202	***
School Activity	<---	Work Status	-1.193	.895	-1.333	.182
School Activity	<---	High Income $75K to $100K	2.834	1.139	2.489	.013
School Activity	<---	Middle Income $30K to $75K	8.821	.944	9.343	***
School Activity	<---	Low Income <30K	2.193	1.485	1.477	.140
School Activity	<---	Number of Children	.538	.837	.643	.520
School Activity	<---	Household Size	2.551	.670	3.806	***
School Activity	<---	Number of Vehicles	.844	.521	1.619	.105
Travel Time to Maintenance	<---	Number of Vehicles	-.772	.187	-4.119	***
Travel Time to Maintenance	<---	Household Size	-2.551	.241	-10.596	***
Travel Time to Maintenance	<---	Number of Children	1.981	.301	6.589	***
Travel Time to Maintenance	<---	Low Income <30K	-.598	.533	-1.121	.262
Travel Time to Maintenance	<---	Middle Income $30K to $75K	3.220	.339	9.493	***
Travel Time to Maintenance	<---	High Income $75K to $100K	-1.044	.409	-2.553	.011
Travel Time to Maintenance	<---	Work Status	-1.166	.321	-3.630	***
Travel Time to Maintenance	<---	Student Status	-3.347	.478	-7.007	***
Travel Time to Maintenance	<---	College Educated	-4.401	.264	-16.664	***
Travel Time to Maintenance	<---	Female	-1.886	.266	-7.080	***
Travel Time to Maintenance	<---	Young <35 years	1.704	.433	3.936	***
Travel Time to Maintenance	<---	Middle Age 35 to 55 years	4.081	.356	11.458	***
Travel Time to Maintenance	<---	Weekday	4.406	.285	15.448	***
Travel Time to Discretionary	<---	Number of Vehicles	.839	.140	6.012	***
Travel Time to Discretionary	<---	Household Size	-2.999	.179	-16.727	***
Travel Time to Discretionary	<---	Number of Children	3.112	.224	13.889	***
Travel Time to Discretionary	<---	Low Income <30K	.654	.397	1.646	.100
Travel Time to Discretionary	<---	Middle Income $30K to $75K	.010	.253	.038	.970
Travel Time to Discretionary	<---	High Income $75K to $100K	-1.840	.305	-6.041	***
Travel Time to Discretionary	<---	Work Status	-.403	.239	-1.685	.092
Travel Time to Discretionary	<---	Student Status	-.403	.357	-1.130	.259
Travel Time to Discretionary	<---	College Educated	-2.199	.197	-11.180	***
Travel Time to Discretionary	<---	Female	-1.451	.198	-7.312	***
Travel Time to Discretionary	<---	Young <35 years	2.988	.332	9.004	***
Travel Time to Discretionary	<---	Middle Age 35 to 55 years	1.927	.267	7.209	***
Travel Time to Discretionary	<---	Weekday	.031	.212	.144	.886
Home Activity	<---	Confidence Ellipse Area	-.027	.002	-12.092	***
Work Activity	<---	Confidence Ellipse Area	-.003	.001	-2.201	.028
Maintenance Activity	<---	Confidence Ellipse Area	.006	.001	5.082	***
Discretionary Activity	<---	Confidence Ellipse Area	.014	.001	9.886	***
School Activity	<---	Confidence Ellipse Area	.000	.000	.699	.485
Travel Time to Maintenance	<---	Confidence Ellipse Area	.002	.000	11.542	***
Travel Time to Discretionary	<---	Confidence Ellipse Area	.001	.000	6.009	***
Home Activity	<---	Kernel Density Area	-.043	.287	-.149	.881
Work Activity	<---	Kernel Density Area	-.063	.156	-.405	.686

Table A.13 continued

			Estimate	S.E.	C.R.	P
Maintenance Activity	<---	Kernel Density Area	.906	.158	5.731	***
Discretionary Activity	<---	Kernel Density Area	-.888	.184	-4.816	***
School Activity	<---	Kernel Density Area	-.361	.062	-5.854	***
Travel Time to Maintenance	<---	Kernel Density Area	-.083	.022	-3.750	***
Travel Time to Discretionary	<---	Kernel Density Area	-.139	.017	-8.394	***
Home Activity	<---	Modified Kernel Density area (sq Km)	.037	.083	.443	.658
Work Activity	<---	Modified Kernel Density area (sq Km)	-.386	.045	-8.526	***
Maintenance Activity	<---	Modified Kernel Density area (sq Km)	-.064	.046	-1.398	.162
Discretionary Activity	<---	Modified Kernel Density area (sq Km)	.139	.054	2.597	.009
School Activity	<---	Modified Kernel Density area (sq Km)	.108	.018	6.009	***
Travel Time to Maintenance	<---	Modified Kernel Density area (sq Km)	.089	.006	13.837	***
Travel Time to Discretionary	<---	Modified Kernel Density area (sq Km)	.090	.005	18.634	***
Home Activity	<---	Coefficient of Variation Trips	93.340	10.212	9.140	***
Work Activity	<---	Coefficient of Variation Trips	-42.239	5.546	-7.616	***
Maintenance Activity	<---	Coefficient of Variation Trips	4.533	5.637	.804	.421
Discretionary Activity	<---	Coefficient of Variation Trips	12.515	6.571	1.904	.057
School Activity	<---	Coefficient of Variation Trips	-14.706	2.194	-6.702	***
Travel Time to Maintenance	<---	Coefficient of Variation Trips	-15.720	.790	-19.906	***
Travel Time to Discretionary	<---	Coefficient of Variation Trips	-6.142	.588	-10.443	***
Home Activity	<---	Coefficient of Variation Distance	-15.741	5.123	-3.072	.002
Work Activity	<---	Coefficient of Variation Distance	-54.505	2.785	-19.570	***
Maintenance Activity	<---	Coefficient of Variation Distance	32.850	2.819	11.655	***
Discretionary Activity	<---	Coefficient of Variation Distance	13.883	3.300	4.207	***
School Activity	<---	Coefficient of Variation Distance	3.953	1.104	3.580	***
Travel Time to Maintenance	<---	Coefficient of Variation Distance	3.292	.396	8.304	***
Travel Time to Discretionary	<---	Coefficient of Variation Distance	2.365	.295	8.006	***
Number of Trips	<---	Home Activity	-.003	.000	-24.590	***
Number of Trips	<---	Work Activity	-.002	.000	-14.731	***
Number of Trips	<---	Maintenance Activity	-.003	.000	-21.448	***
Number of Trips	<---	Discretionary Activity	-.004	.000	-31.546	***
Number of Trips	<---	School Activity	-.001	.000	-7.041	***
Number of Trips	<---	Travel Time to Maintenance	.053	.000	114.621	***
Number of Trips	<---	Travel Time to Discretionary	.029	.001	46.800	***

Table A.14 Total Effects of Activity Participation Model with Variability-seeking Nature and Activity Space

	VariabilitySeekingNature	Weekday	Middle Age 35 to 55 years	Young <35 years	Female	College Educated	Student Status	Work Status	High Income $75K to $100K	Middle Income $30K to $75K	Low Income <30K	Number of Vehicles	Number of Children	Household Size	Coeff. Var. Distance	Coeff. Var. Trips	Modified Kernel Density area	Kernel Density Area	Confidence Ellipse Area	Travel Discretionary	Travel Maintenance	School	Discretionary	Maintenance	Work	Home
Coeff. Var. Distance	0.19	0.00	-0.13	-0.19	0.02	0.09	0.08	-0.14	0.07	-0.05	0.08	-0.07	0.07	-0.06	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Coeff. Var. Trips	0.19	0.00	0.00	-0.04	0.02	-0.01	0.00	-0.06	0.02	0.05	0.13	-0.03	0.07	-0.07	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Modified Kernel Density area	0.00	0.00	-7.22	0.00	3.54	-9.86	-1.70	17.31	-13.79	-27.56	-38.46	9.55	-25.06	24.16	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Kernel Density Area	0.00	0.00	-3.40	1.01	1.62	-2.03	-2.21	3.55	-5.97	-10.07	-12.42	0.62	-9.21	7.50	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Confidence Ellipse Area	0.00	0.00	450.08	281.20	-126.41	-290.44	-260.06	59.80	-27.59	19.72	152.11	301.39	-23.92	-2.11	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Travel Discretionary	-0.70	0.03	1.80	2.85	-1.53	-2.74	-0.26	0.72	-2.21	-1.48	-1.59	1.84	1.86	-1.59	2.37	-6.14	0.09	-0.14	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Travel Maintenance	-2.31	4.41	4.22	2.14	-2.20	-5.21	-3.58	0.62	-1.86	0.66	-4.54	0.81	-0.40	-0.12	3.29	-15.72	0.09	-0.08	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
School	-1.99	2.52	-1.75	-7.93	-0.46	7.63	19.77	-0.32	3.52	8.54	0.92	1.87	0.41	3.24	3.95	-14.71	0.11	-0.36	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Discretionary	4.90	-29.25	67.39	112.60	-41.47	-30.64	-25.39	-1.38	-26.60	22.87	21.92	38.17	14.55	-30.51	13.88	12.52	0.14	-0.89	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Maintenance	6.93	0.35	4.24	-3.59	14.24	-18.20	-13.64	-22.80	12.48	2.19	-29.01	-7.26	-10.74	7.84	32.85	4.53	-0.06	0.91	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Work	-17.94	138.68	93.75	85.76	-3.87	4.15	-30.02	110.05	-12.39	24.24	2.86	4.33	-5.42	-11.44	-54.51	-42.24	-0.39	-0.06	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Home	14.39	-128.77	-164.40	-186.33	31.73	36.17	32.90	-101.50	28.29	-45.93	33.78	-36.85	0.51	31.82	-15.74	93.34	0.04	-0.04	-0.03	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Number of Trips	-0.183	0.434	0.289	0.143	-0.122	-0.309	-0.128	0.192	-0.155	-0.03	-0.39	0.062	0.015	-0.028	0.242	-1.227	0.007	-0.007	0	0.029	0.053	-0.001	-0.004	-0.003	-0.002	-0.003

Table A.15 Indirect Effects of Activity Participation Model with Variability-seeking Nature and Activity Space

	VariabilitySeekingNature	Weekday	Middle Age 35 to 55 years	Young <35 years	Female	College Educated	Student Status	Work Status	High Income $75K to $100K	Middle Income $30K to $75K	Low Income <30K	Number of Vehicles	Number of Children	Household Size	Coeff. Var. Distance	Coeff. Var. Trips	Modified Kernel Density area	Kernel Density Area	Confidence Ellipse Area	Travel Discretionary	Travel Maintenance	School	Discretionary	Maintenance	Work	Home
Coeff. Var. Distance	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Coeff. Var. Trips	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Modified Kernel Density area	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Kernel Density Area	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Confidence Ellipse Area	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Travel Discretionary	-0.701	0	-0.128	-0.137	-0.083	-0.541	0.141	1.118	-0.37	-1.488	-2.243	0.998	-1.249	1.413	0	0	0	0	0	0	0	0	0	0	0	0
Travel Maintenance	-2.305	0	0.134	0.438	-0.316	-0.808	-0.231	1.787	-0.818	-2.556	-3.946	1.583	-2.376	2.433	0	0	0	0	0	0	0	0	0	0	0	0
School	-1.994	0	0.119	-0.461	-0.464	0.093	0.831	0.877	0.688	-0.281	-1.271	1.024	-0.13	0.692	0	0	0	0	0	0	0	0	0	0	0	0
Discretionary	4.895	0	6.413	0.012	-2.232	-2.53	-0.761	-2.526	4.101	5.376	10.557	3.701	6.145	-5	0	0	0	0	0	0	0	0	0	0	0	0
Maintenance	6.932	0	-4.239	-3.59	1.118	-0.09	-0.802	-2.334	-2.457	-8.544	-4.668	-0.589	-4.357	3.013	0	0	0	0	0	0	0	0	0	0	0	0
Work	-17.94	0	9.21	10.807	-2.894	0.292	-3.019	2.838	1.532	11.603	5.328	0.361	3.76	-3.707	0	0	0	0	0	0	0	0	0	0	0	0
Home	14.39	0	-10.278	-8.03	4.962	4.959	5.771	-4.115	0.899	4.479	6.312	-9.065	5.448	-4.956	0	0	0	0	0	0	0	0	0	0	0	0
Number of Trips	-0.183	0.434	0.289	0.143	-0.122	-0.309	-0.128	0.192	-0.155	-0.03	-0.39	0.062	0.015	-0.028	0.242	-1.227	0.007	-0.007	0	0	0	0	0	0	0	0

Table A.16 Estimates of Activity Participation Model with Modified Kernel Density Activity Space

			Estimate	S.E.	C.R.	P
ModKDArea	<---	totveh	9.549	.288	33.152	***
ModKDArea	<---	hhsize	24.164	.366	66.007	***
ModKDArea	<---	numchildren	-25.058	.454	-55.144	***
ModKDArea	<---	LowIncome	-38.463	.827	-46.500	***
ModKDArea	<---	MidIncome	-27.555	.514	-53.623	***
ModKDArea	<---	HighIncome	-13.792	.651	-21.202	***
ModKDArea	<---	WorkStatus	17.309	.496	34.912	***
ModKDArea	<---	studentStatus	-1.704	.723	-2.358	.018
ModKDArea	<---	CollegeStatus	-9.860	.416	-23.718	***
ModKDArea	<---	Female	3.541	.426	8.320	***
ModKDArea	<---	MiddleAge	-7.219	.475	-15.214	***
ActHome	<---	totveh	-32.449	1.942	-16.710	***
ActHome	<---	hhsize	33.530	1.655	20.262	***
ActHome	<---	LowIncome	24.757	5.784	4.280	***
ActHome	<---	MidIncome	-50.558	3.921	-12.894	***
ActHome	<---	HighIncome	25.402	5.138	4.944	***
ActHome	<---	WorkStatus	-100.199	3.218	-31.135	***
ActHome	<---	studentStatus	33.749	5.830	5.789	***
ActHome	<---	CollegeStatus	35.022	3.122	11.219	***
ActHome	<---	Female	30.390	3.125	9.724	***
ActHome	<---	YoungAge	-187.726	4.871	-38.541	***
ActHome	<---	MiddleAge	-166.813	4.091	-40.777	***
ActHome	<---	WeekDay	-128.357	3.089	-41.554	***
ActWork	<---	totveh	8.126	1.230	6.605	***
ActWork	<---	hhsize	-2.730	1.551	-1.761	.078
ActWork	<---	numchildren	-14.147	1.882	-7.518	***
ActWork	<---	MidIncome	17.210	2.069	8.317	***
ActWork	<---	HighIncome	-14.951	2.718	-5.501	***
ActWork	<---	WorkStatus	117.042	2.208	52.997	***
ActWork	<---	studentStatus	-31.159	3.344	-9.319	***
ActWork	<---	YoungAge	83.546	3.062	27.281	***
ActWork	<---	MiddleAge	90.871	2.427	37.439	***
ActWork	<---	WeekDay	138.889	1.992	69.738	***
ActMaintenance	<---	MiddleAge	7.077	2.101	3.368	***
ActMaintenance	<---	Female	12.914	1.872	6.900	***
ActMaintenance	<---	CollegeStatus	-16.538	1.836	-9.007	***
ActMaintenance	<---	studentStatus	-14.457	3.209	-4.505	***
ActMaintenance	<---	WorkStatus	-26.451	2.223	-11.897	***
ActMaintenance	<---	HighIncome	14.317	2.894	4.947	***
ActMaintenance	<---	MidIncome	6.016	2.318	2.595	.009
ActMaintenance	<---	LowIncome	-24.933	3.685	-6.765	***
ActMaintenance	<---	numchildren	-7.898	1.798	-4.392	***
ActMaintenance	<---	hhsize	4.329	1.530	2.829	.005
ActMaintenance	<---	totveh	-9.237	1.243	-7.429	***

Table A.16 continued

			Estimate	S.E.	C.R.	P
ActDiscretionary	<---	totveh	37.523	1.422	26.388	***
ActDiscretionary	<---	hhsize	-29.546	1.692	-17.467	***
ActDiscretionary	<---	numchildren	12.846	1.941	6.618	***
ActDiscretionary	<---	LowIncome	18.352	4.123	4.451	***
ActDiscretionary	<---	MidIncome	21.606	2.584	8.362	***
ActDiscretionary	<---	HighIncome	-27.211	3.338	-8.151	***
ActDiscretionary	<---	studentStatus	-24.946	3.777	-6.604	***
ActDiscretionary	<---	CollegeStatus	-30.502	2.137	-14.270	***
ActDiscretionary	<---	Female	-41.819	2.187	-19.122	***
ActDiscretionary	<---	YoungAge	112.651	3.563	31.621	***
ActDiscretionary	<---	MiddleAge	67.237	2.643	25.440	***
ActDiscretionary	<---	WeekDay	-29.450	2.274	-12.950	***
ActSchool	<---	WeekDay	2.518	.795	3.169	.002
ActSchool	<---	MiddleAge	-1.908	.857	-2.227	.026
ActSchool	<---	YoungAge	-8.263	1.185	-6.971	***
ActSchool	<---	CollegeStatus	7.569	.718	10.538	***
ActSchool	<---	studentStatus	19.651	1.318	14.907	***
ActSchool	<---	HighIncome	3.058	1.070	2.859	.004
ActSchool	<---	MidIncome	8.243	.802	10.276	***
ActSchool	<---	hhsize	3.806	.322	11.805	***
TripMaintenance	<---	totveh	-.139	.179	-.781	.435
TripMaintenance	<---	hhsize	-2.311	.229	-10.097	***
TripMaintenance	<---	numchildren	1.819	.280	6.487	***
TripMaintenance	<---	LowIncome	-1.435	.497	-2.890	.004
TripMaintenance	<---	MidIncome	3.070	.310	9.891	***
TripMaintenance	<---	HighIncome	-.675	.402	-1.679	.093
TripMaintenance	<---	WorkStatus	-1.024	.317	-3.232	.001
TripMaintenance	<---	studentStatus	-3.395	.450	-7.544	***
TripMaintenance	<---	CollegeStatus	-4.293	.260	-16.503	***
TripMaintenance	<---	Female	-2.571	.265	-9.694	***
TripMaintenance	<---	YoungAge	2.299	.426	5.398	***
TripMaintenance	<---	MiddleAge	4.949	.344	14.367	***
TripMaintenance	<---	WeekDay	4.385	.263	16.648	***
TripDiscretionary	<---	totveh	1.159	.129	8.969	***
TripDiscretionary	<---	hhsize	-3.102	.166	-18.730	***
TripDiscretionary	<---	numchildren	3.371	.205	16.458	***
TripDiscretionary	<---	HighIncome	-1.503	.261	-5.762	***
TripDiscretionary	<---	WorkStatus	-.434	.226	-1.920	.055
TripDiscretionary	<---	CollegeStatus	-2.115	.193	-10.946	***
TripDiscretionary	<---	Female	-1.778	.197	-9.021	***
TripDiscretionary	<---	YoungAge	2.944	.309	9.519	***
TripDiscretionary	<---	MiddleAge	2.270	.253	8.967	***
ActHome	<---	ModKDArea	-.222	.032	-7.033	***
ActWork	<---	ModKDArea	-.337	.017	-19.359	***
ActMaintenance	<---	ModKDArea	.163	.018	9.091	***
ActDiscretionary	<---	ModKDArea	.002	.021	.078	.938

Table A.16 continued

			Estimate	S.E.	C.R.	P
ActSchool	<---	ModKDArea	.033	.006	5.191	***
TripMaintenance	<---	ModKDArea	.093	.003	36.814	***
TripDiscretionary	<---	ModKDArea	.064	.002	34.989	***
numtrips	<---	ActHome	-.003	.000	-24.590	***
numtrips	<---	ActWork	-.002	.000	-14.733	***
numtrips	<---	ActMaintenance	-.003	.000	-21.449	***
numtrips	<---	ActDiscretionary	-.004	.000	-31.544	***
numtrips	<---	ActSchool	-.001	.000	-7.038	***
numtrips	<---	TripMaintenance	.053	.000	114.623	***
numtrips	<---	TripDiscretionary	.029	.001	46.800	***

Table A.17 Total Effects of Activity Participation Model with Modified Kernel Density Activity Space

	WeekDay	Middle Age 35 to 55 years	Young <35 years	Female	College Educated	Student Status	Work Status	High Income $75K to $100K	Middle Income $30K to $75K	Low Income <30K	Number of Vehicles	Number of Children	Household Size	Modified Kernel Density area	Travel Discretionary	Travel Maintenance	School	Discretionary	Maintenance	Work	Home
Modified Kernel Density area	0.00	-7.22	0.00	3.54	-9.86	-1.70	17.31	-13.79	-27.56	-38.46	9.55	-25.06	24.16	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Travel Discretionary	0.00	1.81	2.94	-1.55	-2.75	-0.11	0.67	-2.39	-1.76	-2.46	1.77	1.77	-1.56	0.06	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Travel Maintenance	4.39	4.28	2.30	-2.24	-5.21	-3.55	0.59	-1.96	0.51	-5.01	0.75	-0.51	-0.07	0.09	0.00	0.00	0.00	0.00	0.00	0.00	0.00
School	2.52	-2.15	-8.26	0.12	7.25	19.60	0.57	2.61	7.34	-1.26	0.31	-0.82	4.60	0.03	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Discretionary	-29.45	67.23	112.65	-41.81	-30.52	-24.95	0.03	-27.23	21.56	18.29	37.54	12.81	-29.51	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Maintenance	0.00	5.90	0.00	13.49	-18.15	-14.74	-23.63	12.07	1.52	-31.21	-7.68	-11.99	8.27	0.16	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Work	138.89	93.30	83.55	-1.19	3.32	-30.59	111.21	-10.30	26.49	12.96	4.91	-5.70	-10.87	-0.34	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Home	-128.36	-165.21	-187.73	29.60	37.21	34.13	-104.05	28.47	-44.44	33.30	-34.57	5.57	28.16	-0.22	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Number of Trips	0.43	0.29	0.15	-0.12	-0.31	-0.12	0.19	-0.17	-0.05	-0.44	0.06	0.00	-0.02	0.01	0.03	0.05	0.00	0.00	0.00	0.00	0.00

Table A.18 Indirect Effects of Activity Participation Model with Modified Kernel Density Activity Space

	WeekDay	Middle Age 35 to 55 years	Young <35 years	Female	College Educated	Student Status	Work Status	High Income $75K to $100K	Middle Income $30K to $75K	Low Income <30K	Number of Vehicles	Number of Children	Household Size	Modified Kernel Density area	Travel Discretionary	Travel Maintenance	School	Discretionary	Maintenance	Work	Home
Modified Kernel Density area	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Travel Discretionary	0	-0.462	0	0.227	-0.631	-0.109	1.108	-0.883	-1.763	-2.461	0.611	-1.604	1.546	0	0	0	0	0	0	0	0
Travel Maintenance	0	-0.671	0	0.329	-0.917	-0.158	1.609	-1.282	-2.562	-3.576	0.888	-2.33	2.247	0	0	0	0	0	0	0	0
School	0	-0.237	0	0.116	-0.324	-0.056	0.568	-0.453	-0.904	-1.262	0.313	-0.822	0.793	0	0	0	0	0	0	0	0
Discretionary	0	-0.012	0	0.006	-0.016	-0.003	0.028	-0.022	-0.044	-0.062	0.015	-0.04	0.039	0	0	0	0	0	0	0	0
Maintenance	0	-1.178	0	0.578	-1.609	-0.278	2.824	-2.25	-4.495	-6.275	1.558	-4.088	3.942	0	0	0	0	0	0	0	0
Work	0	2.432	0	-1.193	3.322	0.574	-5.832	4.647	9.284	12.959	-3.217	8.443	-8.141	0	0	0	0	0	0	0	0
Home	0	1.604	0	-0.787	2.191	0.379	-3.846	3.065	6.123	8.547	-2.122	5.568	-5.37	0	0	0	0	0	0	0	0
Number of Trips	0.43	0.292	0.15	-0.122	-0.311	-0.123	0.189	-0.165	-0.046	-0.435	0.055	0.004	-0.022	0.008	0	0	0	0	0	0	0

REFERENCES

[1] R. M. Pendyala and E. I. PAS, Multi-Day and Multi-Period Data for Travel Demand Analysis and Modeling Transportation Research Board, 2000.

[2] P. Jones and M. Clarke, "Significance and Measurement of Variability in Travel Behavior: A Discussion Paper," 1987.

[3] H. L. Stefan Schonfelder, Randall Guensler, Jennifer Ogle, Kay W. Axhausen. (2006, Analysis of Commute Atlanta Instrumented Vehicle GPS Data: Destination Choice Behavior and Activity.

[4] E. I. Pas and F. S. Koppelman, "An examination of the determinants of day-to-day variability in individuals' urban travel behavior," Transportation, vol. 14, pp. 3-20, 1987.

[5] C. R. Bhat and F. S. Koppelman, "Activity-based modeling of travel demand," in Handbook of transportation Science, ed: Springer, 1999, pp. 35-61.

[6] Atlanta Regional Commission. (2014). Travel Demand Model. Available: http://www.atlantaregional.com/transportation/travel-demand-model

[7] Atlanta Regional Commission, "Activity-Based Travel Demand Model Specifications: Coordinated Travel - Regional Activity Based Modeling Platform (CT-RAMP) for the Atlanta Region," Atlanta Regional Commission, Atlanta2012.

[8] San Dieog Regional Planning Agency. (November 07, 2014). Activity Based Modeling - San Diego Regional Planning Agency. Available: http://www.sandag.org/index.asp?subclassid=120&fuseaction=home.subclasshome#ABMSteps

[9] NuStats, "2001 Atlanta Household Travel Survey," Atlanta Regional CommissionApril 2003.

[10] NuStats, "Regional Travel Survey - Final Report," Atlanta Regional CommissionNovember 2011.

[11] Y. Xu, "Effective GPS-based panel survey sample size for urban travel behavior studies," PhD, Department of Civil and Environmental Engineering, Georgia Institute of Technology, 2010.

[12] V. Elango and R. Guensler, "An Automated Activity Identification Method for Passively Collected GPS Data," presented at the The 3rd Conference on Innovations in Travel Modeling, Phoenix AZ, 2010.

[13] S. Hanson and J. O. Huff, "Assessing day-to-day variability in complex travel patterns," Transportation Research Record, vol. 891, pp. 18-24, 1981.

[14] R. Schlich and K. W. Axhausen, "Habitual Travel Behaviour: Evidence from a Six-Week Travel Diary " in Transportation (Netherlands) vol. 30, ed: Kluwer Academic Publishers, Martinus Nijhoff Publishers, 2003, pp. p. 13-36.

[15] R. M. Pendyala, L. P. Kostyniuk, and K. G. Goulias, "A repeated cross-sectional evaluation of car ownership," Transportation, vol. 22, pp. 165-184, 1995.

[16] S. Schönfelder, K. W. Axhausen, N. Antille, and M. Bierlaire, Exploring the Potentials of Automatically Collected GPS Data for Travel Behaviour Analysis: A Swedish Data Source: Citeseer, 2002.

[17] J. Ogle, R. Guensler, and V. Elango, "Georgia's commute Atlanta value pricing program: recruitment methods and travel diary response rates," Transportation Research Record: Journal of the Transportation Research Board, vol. 1931, pp. 28-37, 2005.

[18] S. Greaves, S. Fifer, R. Ellison, and G. Germanos, "Development of a global positioning system web-based prompted recall solution for longitudinal travel surveys," Transportation Research Record: Journal of the Transportation Research Board, vol. 2183, pp. 69-77, 2010.

[19] Y. O. Susilo and R. Kitamura, "Analysis of day-to-day variability in an individual's action space: exploration of 6-week Mobidrive travel diary data," Transportation Research Record: Journal of the Transportation Research Board, vol. 1902, pp. 124-133, 2005.

[20] E. I. Pas and S. Sundar, "Intrapersonal variability in daily urban travel behavior: some additional evidence," Transportation, vol. 22, pp. 135-150, 1995.

[21] P. Stopher, E. Clifford, and M. Montes, "Variability of travel over multiple days: analysis of three panel waves," Transportation Research Record: Journal of the Transportation Research Board, vol. 2054, pp. 56-63, 2008.

[22] S. G. Hatcher and H. S. Mahmassani, "Daily variability of route and trip scheduling decisions for the evening commute," Transportation Research Record, 1992.

[23] H. Li, "Investigating morning commute route choice behavior using global positioning systems and multi-day travel data," 2004.

[24] E. I. Pas, "Intrapersonal variability and model goodness-of-fit," Transportation Research Part A: General, vol. 21, pp. 431-438, 1987.

[25] H. Kang and D. M. Scott, "Exploring day-to-day variability in time use for household members," Transportation Research Part A: Policy and Practice, vol. 44, pp. 609-619, 2010.

[26] J. Wolf, R. Guensler, and W. Bachman, "Elimination of the travel diary: Experiment to derive trip purpose from global positioning system travel data," Transportation Research Record: Journal of the Transportation Research Board, vol. 1768, pp. 125-134, 2001.

[27] Y. Lu, S. Zhu, and L. Zhang, "Imputing Trip Purpose Based on GPS Travel Survey Data and Machine Learning Methods," in Transportation Research Board 92nd Annual Meeting, 2013.

[28] S. S. Srivastava G, On the Temporal Variation of Activity Spaces: ETH, Eidgenössische Technische Hochschule Zürich, Institut für Verkehrsplanung und Transportsysteme, July 2003.

[29] ESRI. (2009). ArcGIS Desktop Help - Directional Distribution (Standard Deviational Ellipse). Available: http://webhelp.esri.com/arcgisdesktop/9.2/index.cfm?TopicName=How%20Directional%20Distribution:%20Standard%20Deviational%20Ellipse%20(Spatial%20Statistics)%20works

[30] K. W. Axhausen and T. Gärling, "Activity-based approaches to travel analysis: conceptual frameworks, models, and research problems," Transport Reviews, vol. 12, pp. 323-341, 1992.

[31] P. M. Jones, F. Koppelman, and J.-P. Orfeuil, "Activity analysis: State-of-the-art and future directions," Developments in dynamic and activity-based approaches to travel analysis, pp. 34-55, 1990.

[32] C. R. Bhat, "Recent methodological advances relevant to activity and travel behavior analysis," in International Association of Travel Behavior Research Conference, Austin, Texas, 1997.

[33] D. Prelec, "Matching, maximizing, and the hyperbolic reinforcement feedback function," Psychological Review, vol. 89, p. 189, 1982.

[34] D. Kahneman and A. Tversky, "Prospect theory: An analysis of decision under risk," Econometrica: Journal of the Econometric Society, pp. 263-291, 1979.

[35] B. D. Meyer, "Unemployment insurance and unemployment spells," ed: National Bureau of Economic Research Cambridge, Mass., USA, 1991.

[36] T. F. Golob, "Structural equation modeling for travel behavior research," Transportation Research Part B: Methodological, vol. 37, pp. 1-25, 2003.

[37] A. R. Kuppam and R. M. Pendyala, "A structural equations analysis of commuters' activity and travel patterns," Transportation, vol. 28, pp. 33-54, 2001.

[38] J.-H. Chung and Y. Ahn, "Structural equation models of day-to-day activity participation and travel behavior in a developing country," Transportation Research Record: Journal of the Transportation Research Board, vol. 1807, pp. 109-118, 2002.

[39] G. D. Garson. (2012, 05-14-14). Structural Equation Modeling [Electronic].

[40] T. F. Golob and M. G. McNally, "A model of activity participation and travel interactions between household heads," Transportation Research Part B: Methodological, vol. 31, pp. 177-194, 1997.

[41] T. F. Golob, "A model of household demand for activity participation and mobility," 1996.

[42] X. Lu and E. I. Pas, "Socio-demographics, activity participation and travel behavior," Transportation Research Part A: Policy and Practice, vol. 33, pp. 1-18, 1999.

[43] J.-H. Chung, S. Kim, Y.-K. Lee, and Y.-S. Choi, "Multilevel structural equation model for activity participation and travel behavior: Data from the Puget sound transportation panel," Transportation Research Record: Journal of the Transportation Research Board, vol. 1898, pp. 52-60, 2004.

[44] H. Kang and D. M. Scott, "Joint or Solo: Structural Equations Model of Household Activity Time Allocation Patterns," in Transportation Research Board 87th Annual Meeting, 2008.

[45] C. Weis and K. W. Axhausen, "Induced travel demand: Evidence from a pseudo panel data based structural equations model," Research in Transportation Economics, vol. 25, pp. 8-18, 2009.

[46] D. Wang and T. Lin, "Built environments, social environments, and activity-travel behavior: a case study of Hong Kong," Journal of Transport Geography, vol. 31, pp. 286-295, 2013.

[47] J. de Abreu e Silva, C. Morency, and K. G. Goulias, "Using structural equations modeling to unravel the influence of land use patterns on travel behavior of workers in Montreal," Transportation Research Part A: Policy and Practice, vol. 46, pp. 1252-1264, 2012.

[48] Y. Xu, L. I. Zuyeva, D. Kall, V. V. Elango, and R. Guensler, Mileage-Based Value Pricing: Phase II Case Study Implications of Commute Atlanta Project: Transportation Research Board, 2009.

[49] G. R. Jun J, Ogle J, "Smoothing Methods to Minimize Impact of Global Positiong System Random Error on Travel Distance, Speed, and Acceleration Profile Estimates," Transportation Research Record: Journal of the Transportation Research Board, vol. 1972, pp. 141-150, 2006.

[50] M. Lee-Gosselin, P. Bonnel, J. Zmud, J. Madre, Transport Survey Methods: Keeping Up With the Changing World. UK: Emerald Group Publishing Limited, 2009.

[51] B. C. Bricka Stacey, "Comapative Analysis of Global Positioning System-Based and Travel Survey-Based Data," Transportation Research Record: Journal of the Transportation Research Board, vol. 1972, pp. 9-20, 2006 2006.

[52] L. M. Wolf J, Meyers J and Arce C, "Trip Rate Analysis in GPS-Enhanced Personal Travel Surveys," in International Conference on Transport Survey Quality and Innovation, Kruger Park, South Africa, 2001.

[53] D. Peason, "Global Positioning System (GPS) and travel survey results: Results from the 1997 Austin Household Survey," in Eigth Conference on the Application of Transportation Planning Methods, Corpus Christi, Tx, April 2001.

[54] B. Pierce, J. Casas, and G. Giaimo, "Estimating trip rate under-reporting: preliminary results from the Ohio household travel survey," in Transportation Research Board 82nd Annual Meeting Pre-print CD–ROM, 2003.

[55] K. Axhausen, S. Schönfelder, J. Wolf, M. Oliveira, and U. Samaga, "Eighty weeks of gps traces, approaches to enriching trip information," in Transportation Research Board 83rd Annual Meeting Pre-print CD-ROM, 2004.

[56] P. Stopher, E. Clifford, J. Zhang, and C. FitzGerald, Deducing mode and purpose from GPS data: Institute of Transport and Logistics Studies, 2008.

[57] C. M. U. Cosma Shalizi. (04/17/2013). Lecture 22 - Classification and Regression Trees, . Available: http://www.stat.cmu.edu/~cshalizi/350/lectures/22/lecture-22.pdf

[58] A. Santos, N. McGuckin, H. Y. Nakamoto, D. Gray, and S. Liss, "Summary of travel trends: 2009 national household travel survey," 2011.

[59] V. V. Elango, R. Guensler, and J. H. Ogle, Day-to-Day Travel Variability in the Commute Atlanta, Georgia, Study: Transportation Research Record; Journal of Transportation Research Board, 2007.

[60] R. Schlich, S. Schönfelder, S. Hanson, and K. W. Axhausen, "Structures of leisure travel: temporal and spatial variability," Transport Reviews, vol. 24, pp. 219-237, 2004.

[61] P. Stopher, Collecting, managing, and assessing data using sample surveys: Cambridge University Press, 2012.

[62] S. Manikandan, "Measures of dispersion," Journal of pharmacology & pharmacotherapeutics, vol. 2, p. 315, 2011.

[63] P. R. Stopher, K. Kockelman, S. P. Greaves, and E. Clifford, "Reducing burden and sample sizes in multiday household travel surveys," Transportation Research Record: Journal of the Transportation Research Board, vol. 2064, pp. 12-18, 2008.

[64] E. Spissu, I. Meloni, and B. Sanjust, "Behavioral analysis of choice of daily route with data from global positioning system," Transportation Research Record: Journal of the Transportation Research Board, vol. 2230, pp. 96-103, 2011.

[65] B. Efron and R. Tibshirani, An introduction to the bootstrap vol. 57: CRC press, 1993.

[66] J. Felsenstein, "Confidence limits on phylogenies: an approach using the bootstrap," Evolution, pp. 783-791, 1985.

[67] M. Eichler. (Oct 2013). Introduction to Bootstrap. Available: http://galton.uchicago.edu/~eichler/stat24600/Handouts/bootstrap.pdf

[68] P. E. McKnight and J. Najab, "Mann-Whitney U Test," Corsini Encyclopedia of Psychology, 2010.

[69] H. B. Mann and D. R. Whitney, "On a test of whether one of two random variables is stochastically larger than the other," The annals of mathematical statistics, vol. 18, pp. 50-60, 1947.

[70] M. L. Marx and R. J. Larsen, Introduction to mathematical statistics and its applications: Pearson/Prentice Hall, 2006.

[71] E. Fieller, H. Hartley, and E. Pearson, "Tests for rank correlation coefficients. I," Biometrika, vol. 44, pp. 470-481, 1957.

[72] K. W. Axhausen, S. Schonfelder, "Measuring the Size and Structure of Human Activity Spaces - The Longitudinal Perspective," presented at the ETC, 2002.

[73] H. Li, R. Guensler, and J. Ogle, "Impact of Objective Route Attributes on the Choice of Primary Morning Commute Route," in Transportation Research Board 85th Annual Meeting, 2006.

[74] S. Schönfelder, "Urban rhythms: Modelling the rhythms of individual travel behaviour," 2006.

[75] S. Schönfelder and K. W. Axhausen, Urban rhythms and travel behaviour: spatial and temporal phenomena of daily travel: Ashgate Publishing, Ltd., 2010.

[76] M. B. R.K. Rai, Marcel Rieser, V.S. Vaze, Stefan Schonfelder, Kay W. Axhausen, "Capturing Human Activity Spaces: New Geometries," in Transportation Research Board 2007 Annual Meeting, Washington DC, 2007.

[77] M.-P. Kwan, "Interactive geovisualization of activity-travel patterns using three-dimensional geographical information systems: a methodological exploration with a large data set," Transportation Research Part C: Emerging Technologies, vol. 8, pp. 185-203, 2000.

[78] V. V. Elango, S. Khoeini, Y. Xu, and R. Guensler, "Longitudinal GPS Travel Data and Breach of Privacy via Enhanced Spatial and Demographic Analysis," in Transportation Research Board 92nd Annual Meeting, 2013.

[79] Y. Yang and A. V. Diez-Roux, "Walking distance by trip purpose and population subgroups," American journal of preventive medicine, vol. 43, pp. 11-19, 2012.

[80] L. Aultman-Hall, M. Roorda, and B. W. Baetz, "Using GIS for evaluation of neighborhood pedestrian accessibility," Journal of Urban Planning and Development, vol. 123, pp. 10-17, 1997.

[81] F. Atash, "Redesigning suburbia for walking and transit: emerging concepts," Journal of Urban Planning and Development, vol. 120, pp. 48-57, 1994.

[82] J. Wolf, R. Guensler, S. Washington, and W. Bachman, "High-emitting vehicle characterization using regression tree analysis," Transportation Research Record: Journal of the Transportation Research Board, vol. 1641, pp. 58-65, 1998.

[83] S. L. Hallmark, R. Guensler, and I. Fomunung, "Characterizing on-road variables that affect passenger vehicle modal operation," Transportation Research Part D: Transport and Environment, vol. 7, pp. 81-98, 2002.

[84] I. Fomunung, S. Washington, R. Guensler, and W. Bachman, "Validation of the MEASURE automobile emissions model: a statistical analysis," Journal of Transportation and Statistics (September 2000), pp. 65-84, 2001.

[85] V. V. Elango and R. Guensler, "Vehicle Occupancy Data Collection, Screening, and Evaluation," in Transportation Research Board 93rd Annual Meeting, 2014.

[86] R. E. Schumacker and R. G. Lomax, A beginner's guide to structural equation modeling: Psychology Press, 2004.

[87] Wikipedia. (2014, 08/18/2014). Structural Equation Modeling. Available: http://en.wikipedia.org/wiki/Structural_equation_modeling

[88] B. S. Everitt, An introduction to latent variable models: Springer, 1984.

[89] G. Marcoulides and T. Raykov, "A first course in structural equation modeling," ed: L. Erlbaum Associates, Mahwah (NJ), 2000.

[90] R. B. Kline, Principles and practice of structural equation modeling: Guilford press, 2011.

[91] J. C. Loehlin, Latent variable models: An introduction to factor, path, and structural analysis: Lawrence Erlbaum Associates Publishers, 1998.

[92] J. P. Stevens, Applied multivariate statistics for the social sciences: Routledge, 2012.

[93] P. M. Bentler and C.-P. Chou, "Practical issues in structural modeling," Sociological Methods & Research, vol. 16, pp. 78-117, 1987.

[94] M. W. Browne, R. Cudeck, and K. A. Bollen, "Alternative ways of assessing model fit," Sage Focus Editions, vol. 154, pp. 136-136, 1993.

[95] T. F. Golob, "A simultaneous model of household activity participation and trip chain generation," Transportation Research Part B: Methodological, vol. 34, pp. 355-376, 2000.

[96] J. Bongaarts, "Household size and composition in the developing world in the 1990s," Population studies, vol. 55, pp. 263-279, 2001.

[97] J. Holtzclaw, R. Clear, H. Dittmar, D. Goldstein, and P. Haas, "Location efficiency: Neighborhood and socio-economic characteristics determine auto ownership and use-studies in Chicago, Los Angeles and San Francisco," Transportation Planning and Technology, vol. 25, pp. 1-27, 2002.

[98] D. V. C. Gretchen Livingston. (2010, 08/18/2014). The New Demography of American Motherhood. Available: http://www.pewsocialtrends.org/2010/05/06/the-new-demography-of-american-motherhood/

[99] National Center for Education Statistics". (2014, 08/18/2014). Characteristics of Post Secondary Students. Available: http://nces.ed.gov/programs/coe/indicator_csb.asp

[100] National Center for Education Statistics". (2013, 08/18/2014). Trends in Employment Rates by Educational Attainment Available: http://nces.ed.gov/programs/coe/indicator_tba.asp

[101] Atlanta Regional Commission, "The Travel Forecasting Model Set for the Atlanta Region," Atlanta Regional Commission, Atlanta2011.

[102] J. L. Arbuckle, "AMOS," 22 ed: SPSS, 2013.

[103] J. Nevitt and G. R. Hancock, "Performance of bootstrapping approaches to model test statistics and parameter standard error estimation in structural equation modeling," Structural Equation Modeling, vol. 8, pp. 353-377, 2001.

[104] V. V. Elango and R. Guensler. (2014). Commute Warrior. Available: https://play.google.com/store/apps/details?id=com.gatech.cw3&hl=en